

|  |
|--|
| PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.                                   |
| PROSITE; PS00108; PROTEIN_KINASE_ST; 1.                                    |
| Alternative splicing; ATP-binding; Nuclear protein; Phosphorylation;       |
| KW Serine/threonine-protein kinase; Transferrase; Tyrosine-protein kinase. |
| FTTFT  |
| DOMAIN 159 475 Protein kinase.   |
| NP_BIND 165 173 ATP (By similarity).                                       |
| BINDING 189 189 ATP (By similarity).                                       |
| ACT_SITE 286 286 Proton acceptor (By similarity).                          |
| VARSPLIT 1 180 Missing (in isoform 2).                                     |
| FTTFT  |
| MUTAGEN 189 189 /FTTD=VSP_008205.  |
| CONFLICT 58 58 K->R; Loss of function.                                     |
| CONFLICT 141 141 E -> Q (in Ref. 4).                                       |
| CONFLICT 344 344 D -> N (in Ref. 4).                                       |
| SEQUENCE 481 AA; 57344 MW; F2C5G965900CI2AA CRC64;                         |
| Query Match 93.8%; Score 2261.5; DB 1; Length 481;                         |
| Best Local Similarity 93.5%; Pred No. 5.1e-132;                            |
| Matches 420; Conservative 8; Mismatches 10; Indels 11; Gaps 2;             |
| 8 SH-SVEEDTH-----PSHYLEARSLNERDYDRRYVDYDEYNDYCEGVPRHYHRD 56                |
| 33 SHSSTQENRHCKPHHQPKDSCHYLEARCLNERDYDRRYIDEYNDYCEGVPRHYHRD 92             |
| 57 IESGVRIHCSKSSVRSRRSSPKRKNRHCSSSHOSKSKSHRRKRSRSIDDEEGHLICQSG 116         |
| :   :     :     :     :     :     :     :     :     :     :                |
| 93 VESTYRIHCSKSSVRSRRSSPKRKNRPCASHQSHSKSHRRKRSRSIDDEEGHLICQSG 152          |
| 117 DVLRARYEIVDTLGBGAGKVVECIDHGMDGMHVAVKIVGNVGRYREARSIOVLLEHLN 176         |
| :     :     :     :     :     :     :     :     :     :                    |
| 153 DVLRARYEIVDTLGBGAGKVVECIDHGMDGLHVAKIVGNVGRYREARSIOVLLEHLN 212          |
| 177 STDPNVSFRCVQLMEWFDDHHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRQMAYIQCO 236      |
| 213 STDPNVSFRCVQLMEWFDDHHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRQMAYIQCO 272      |
| 237 SINFLHNKLHTTDLKPENILFKVSDYVVKNSKKRDERTIKNITDKIVDPGSATYDDE 296          |
| 273 SINFLHNKLHTTDLKPENILFKVSDYVVKNSKKRDERTIKNITDKIVDPGSATYDDE 332          |
| 297 HHSTLVSTRHYRAPVIILALGWSQCDVWSIGCILIEYLGLTFVTQTHDSKSHLAMMERI 356        |
| :     :     :     :     :     :     :     :     :     :                    |
| 333 HHSTLVSTRHYRAPVIILALGWSQCDVWSIGCILIEYLGLTFVTQTHDSKSHLAMMERI 392        |
| 357 LGPIPQHMIQTRKEKYFHNNQLDWDESSHASAGYVRRCKPLKEFMLCHDBEHKLFDLV 416         |
| 393 LGPIPAIMIQTRKEKYFHNNQLDWDESSHASAGYVRRCKPLKEFMLCHDBEHKLFDLV 452         |
| 417 RRMLEYDPTORTILDALOHFFDLLKKK 445  |
| :     :     :     :     :     :     :     :     :     :                    |
| 453 RRMLEYDPTARRITLDALOHFFDLLKKK 481                                       |
| RESULT 3   |
| QBNSVB8 PRELIMINARY; PRT; 484 AA.  |
| ID QBNSVB8   |
| AC QBNSVB8   |
| DT 01-OCT-2002 (TreeBLrel. 22, Created)                                    |
| DT 01-OCT-2002 (TreeBLrel. 22, Last sequence update)                       |
| DT 01-WAR-2004 (TreeBLrel. 26, Last annotation update)                     |
| DE CDC-like kinase 1.  |
| OS Name=CLK1;  |
| OS Homo sapiens (Human).   |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;       |
| OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.              |
| NCBI_TaxId=9606;   |
| RN [1]   |
| RN SEQUENCE FROM N.A.  |
| RP TISSUE=Bone;  |
| RC RC  |
| EX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;          |
| RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,                |
| RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,     |
| RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,       |

This Page Blank (uspto)

QY 439 FDLKKK 445  
 Db 421 FDLKKK 427

RESULT 9

US-10-801-671-4  
 ; Sequence 4, Application US/10801671  
 ; Publication No. US20040152123A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: YAN, Chunhua et al.  
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
 ; FILE REFERENCE: THERIOF  
 ; FILE REFERENCE: CLO00758DIV-III  
 ; CURRENT APPLICATION NUMBER: US/10/801,671  
 ; PRIOR FILING DATE: 2004-03-17  
 ; PRIOR APPLICATION NUMBER: 60/227,470  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: 09/810,671  
 ; PRIOR FILING DATE: 2001-03-19  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 427  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-801-671-4

Query Match 95.9%; Score 2312; DB 16; Length 427;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-170;  
 Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 HYLEARSLNERDYDRRYDEYNDYCEGVPRHYHDD 78  
 Db 1 HYLEARSLNERDYDRRYDEYNDYCEGVPRHYHDD 60  
 QY 79 RKRNRHCSHQSKSRSHRRKRSHRSIEDDEGHLCQSGDVLRARYEIVDTLGEAGFGKV 138  
 Db 61 RKRNRHCSHQSKSRSHRRKRSHRSIEDDEGHLCQSGDVLRARYEIVDTLGEAGFGKV 120  
 QY 139 ECIDHGMGMHVAVKVKNVGRYREARSEIOVLEHNLSTDNSVRCVQMLEWFDHGH 198  
 Db 121 ECIDHGMGMHVAVKVKNVGRYREARSEIOVLEHNLSTDNSVRCVQMLEWFDHGH 180  
 QY 199 VCIVFELLGLSYDFIKENSFLPFQIDHIROMAYQICQSFNLFHNKLTHTDLKPNILF 258  
 Db 181 VCIVFELLGLSYDFIKENSFLPFQIDHIROMAYQICQSFNLFHNKLTHTDLKPNILF 240  
 QY 259 VKSDYVVKYNSQKRDERTLKNITDKVDFGSAFYDDEHSTLSTRHYRAPEVILALGW 318  
 Db 241 VKSDYVVKYNSQKRDERTLKNITDKVDFGSAFYDDEHSTLSTRHYRAPEVILALGW 300  
 QY 319 SQPCDWSIGCILIEYLGFTVQTHDSKEHLAMMERILGPPOHMIQKTRKRYFHHNQ 378  
 Db 301 SQPCDWSIGCILIEYLGFTVQTHDSKEHLAMMERILGPPOHMIQKTRKRYFHHNQ 360  
 QY 379 LWDDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRMLEYDPTQRTITLDEALQHPF 438  
 Db 361 LWDDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRMLEYDPTQRTITLDEALQHPF 420  
 QY 439 FDLKKK 445  
 Db 421 FDLKKK 427

RESULT 10

US-09-905-999-25  
 ; Sequence 25, Application US/09905999  
 ; Patent No. US20020106771A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ULIRICH, Axel  
 ; APPLICANT: NAYLER, Oliver

; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS  
 ; FILE REFERENCE: 038602/0431  
 ; CURRENT APPLICATION NUMBER: US/09/905,999  
 ; CURRENT FILING DATE: 2001-07-17  
 ; PRIOR APPLICATION NUMBER: 09/127,248  
 ; PRIOR FILING DATE: 1999-07-31  
 ; PRIOR APPLICATION NUMBER: PCT/IB97/00946  
 ; PRIOR FILING DATE: 1997-06-17  
 ; PRIOR APPLICATION NUMBER: US 08/877,150  
 ; PRIOR FILING DATE: 1997-06-17  
 ; PRIOR APPLICATION NUMBER: US 60/034,286  
 ; PRIOR FILING DATE: 1996-12-19  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 25  
 ; LENGTH: 481  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-905-999-25

Query Match 93.8%; Score 2261.5; DB 9; Length 481;  
 Best Local Similarity 93.5%; Pred. No. 5.2e-166;  
 Matches 420; Conservative 8; Mismatches 10; Indels 11; Gaps 2;  
 QY 8 SH-SVEEDTH-----PSHYLEARSLSNERDYDRRYDEYNDYCEGVPRHYHDD 56  
 Db 33 SHSSTQENRCHCKPHQPFKXSDCHYLEARCLNERDYDRRYDEYNDYCEGVPRHYHDD 92  
 QY 57 IESGYRIHCSKSSVRSRSPKRRHCHSHQSKSRSHRRKRSHRSIEDDEGHLCQSG 116  
 Db 93 VESTYRIHCSKSSVRSRSPKRRHCHSHQSKSRSHRRKRSHRSIEDDEGHLCQSG 152  
 QY 117 DVLRARYEIVDTLGEAGFGKVVECIDHGMGMHVAVKVKNVGRYREARSEIOVLEHNL 176  
 Db 153 DVLRARYEIVDTLGEAGFGKVVECIDHGMGMHVAVKVKNVGRYREARSEIOVLEHNL 212  
 QY 177 STDPSNVRVCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPFQIDHIROMAYQICQ 236  
 Db 213 STDPSNVRVCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPFQIDHIROMAYQICQ 272  
 QY 237 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSQKRDERTLKNITDKVDFGSAFYDDE 296  
 Db 273 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSQKRDERTLKNITDKVDFGSAFYDDE 332  
 QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGFTVQTHDSKEHLAMMERI 356  
 Db 333 HHSTLVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGFTVQTHDSKEHLAMMERI 392  
 QY 357 LGPIPOHMIQKTRKRYFHHNQDDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLV 416  
 Db 393 LGPIPAHMIQKTRKRYFHHNQDDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLV 452  
 QY 417 RRMLEYDPTQRTITLDEALQHPFLLKKK 445  
 Db 453 RRMLEYDPTQRTITLDEALQHPFLLKKK 481

RESULT 11

US-10-267-502-355  
 ; Sequence 355, Application US/10267502  
 ; Publication No. US20040071700A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kim, Jaeseob  
 ; APPLICANT: Galant, Ron  
 ; TITLE OF INVENTION: Obesity Linked Genes  
 ; FILE REFERENCE: LSD-07416  
 ; CURRENT APPLICATION NUMBER: US/10/267,502  
 ; CURRENT FILING DATE: 2003-01-27  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 355  
 ; LENGTH: 481  
 ; TYPE: PRT

This Page Blank (uspto)





This Page Blank (uspto)



**This Page Blank (uspto)**



QY 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSINF 240  
 DB 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSINF 240  
 QY 241 LHNKLTHTDLKPNILFKSDYVVKYNSKMRDERTLKNITDKVVDGSGATYDDEHST 300  
 DB 241 LHNKLTHTDLKPNILFKSDYVVKYNSKMRDERTLKNITDKVVDGSGATYDDEHST 300  
 QY 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360  
 DB 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360  
 QY 361 PQHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420  
 DB 361 PQHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420  
 QY 421 EYDPTQRTITLDEALQHPFFDLKKK 445  
 DB 421 EYDPTQRTITLDEALQHPFFDLKKK 445

RESULT 2

US-10-109-854-2  
 ; Sequence 2, Application US/10109854  
 ; Publication No. US20020119548A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: YAN, Chunhua et al.  
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
 ; TITLE OF INVENTION: THEREOF  
 ; FILE REFERENCE: CL000758DIV  
 ; CURRENT APPLICATION NUMBER: US/10/109,854  
 ; PRIOR FILING DATE: 2002-04-01  
 ; PRIOR APPLICATION NUMBER: 60/227,470  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: 09/810,671  
 ; PRIOR FILING DATE: 2001-03-19  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 445  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-10-109-854-2

Query Match 100.0%; Score 2410; DB 13; Length 445;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-177;  
 Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRDIESTG 60  
 DB 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRDIESTG 60  
 QY 61 YRIHCKSSVRSRSPKRNHCHSHQSRKSHRRKRKRSIEDDEGHLCQSGDVL 120  
 DB 61 YRIHCKSSVRSRSPKRNHCHSHQSRKSHRRKRKRSIEDDEGHLCQSGDVL 120  
 QY 121 ARYEIVDTLGEAGKVECTIDHGMGMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180  
 DB 121 ARYEIVDTLGEAGKVECTIDHGMGMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180  
 QY 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSINF 240  
 DB 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSINF 240  
 QY 241 LHNKLTHTDLKPNILFKSDYVVKYNSKMRDERTLKNITDKVVDGSGATYDDEHST 300  
 DB 241 LHNKLTHTDLKPNILFKSDYVVKYNSKMRDERTLKNITDKVVDGSGATYDDEHST 300  
 QY 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360  
 DB 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360  
 QY 421 EYDPTQRTITLDEALQHPFFDLKKK 445  
 DB 421 EYDPTQRTITLDEALQHPFFDLKKK 445

QY 361 PQHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420  
 DB 361 PQHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420  
 QY 421 EYDPTQRTITLDEALQHPFFDLKKK 445  
 DB 421 EYDPTQRTITLDEALQHPFFDLKKK 445

RESULT 3

US-10-339-656-2  
 ; Sequence 2, Application US/10339656  
 ; Publication No. US20030134319A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: YAN, Chunhua et al.  
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
 ; TITLE OF INVENTION: THEREOF  
 ; FILE REFERENCE: CL000758DIV2  
 ; CURRENT APPLICATION NUMBER: US/10/339,656  
 ; CURRENT FILING DATE: 2003-01-10  
 ; PRIOR APPLICATION NUMBER: 10/109,854  
 ; PRIOR FILING DATE: 2002-04-01  
 ; PRIOR APPLICATION NUMBER: 09/810,671  
 ; PRIOR FILING DATE: 2001-03-19  
 ; PRIOR APPLICATION NUMBER: 60/227,470  
 ; PRIOR FILING DATE: 2000-08-24  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 445  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-10-339-656-2

Query Match 100.0%; Score 2410; DB 14; Length 445;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-177;  
 Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRDIESTG 60  
 DB 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRDIESTG 60  
 QY 61 YRIHCKSSVRSRSPKRNHCHSHQSRKSHRRKRKRSIEDDEGHLCQSGDVL 120  
 DB 61 YRIHCKSSVRSRSPKRNHCHSHQSRKSHRRKRKRSIEDDEGHLCQSGDVL 120  
 QY 121 ARYEIVDTLGEAGKVECTIDHGMGMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180  
 DB 121 ARYEIVDTLGEAGKVECTIDHGMGMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180  
 QY 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSINF 240  
 DB 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSINF 240  
 QY 241 LHNKLTHTDLKPNILFKSDYVVKYNSKMRDERTLKNITDKVVDGSGATYDDEHST 300  
 DB 241 LHNKLTHTDLKPNILFKSDYVVKYNSKMRDERTLKNITDKVVDGSGATYDDEHST 300  
 QY 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360  
 DB 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360  
 QY 361 PQHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420  
 DB 361 PQHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420  
 QY 421 EYDPTQRTITLDEALQHPFFDLKKK 445  
 DB 421 EYDPTQRTITLDEALQHPFFDLKKK 445

RESULT 4

```
US-10-801-671-2
; Sequence 2, Application US/10801671
; Publication No. US20040152123A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV-III
; CURRENT APPLICATION NUMBER: US/10/801,671
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-801-671-2

Query Match      100.0%; Score 2410; DB 16; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.7e-177;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRDRYVDEYRNDYCEGVVPRHYHRDIESG 60
DB 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRDRYVDEYRNDYCEGVVPRHYHRDIESG 60

QY 61 YRIHCKSSVRSRSPKRNHCHSCSHQSRKSHRRKRSRSIEDDEGHLCQSGDVL 120
DB 61 YRIHCKSSVRSRSPKRNHCHSCSHQSRKSHRRKRSRSIEDDEGHLCQSGDVL 120

QY 121 ARYEIVDTLGGAGFKVVECDHGMGDMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180
DB 121 ARYEIVDTLGGAGFKVVECDHGMGDMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180

QY 181 NSVPRCVQMLEFHHGHCIVFELLGLSTYDFIKENSFLPFQIDHROMAYQICQINF 240
DB 181 NSVPRCVQMLEFHHGHCIVFELLGLSTYDFIKENSFLPFQIDHROMAYQICQINF 240

QY 241 LHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDERTLKNNTDIKVDFGSAFYDDEHST 300
DB 241 LHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDERTLKNNTDIKVDFGSAFYDDEHST 300

QY 301 LVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYILGFTVFTQTHDSKEHLAMMERILGPI 360
DB 301 LVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYILGFTVFTQTHDSKEHLAMMERILGPI 360

QY 361 PQHMIQTRKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCHDBEHEKLFDLVRRML 420
DB 361 PQHMIQTRKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCHDBEHEKLFDLVRRML 420

QY 421 EYDPTQRTITLDEALQHPFFDLKKK 445
DB 421 EYDPTQRTITLDEALQHPFFDLKKK 445

RESULT 5
US-10-267-502-352
; Sequence 352, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Patent in version 3.2

; SEQ ID NO 352
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-352

Query Match      96.1%; Score 2315.5; DB 15; Length 481;
Best Local Similarity 96.2%; Pred. No. 3.6e-170;
Matches 432; Conservative 2; Mismatches 4; Indels 11; Gaps 2;

QY 8 SH-SVERDTH-----PSHYLEARSLSNERDYDRDRYVDEYRNDYCEGVVPRHYHRD 56
DB 33 SHSSTQENRHCKPHQFKESDCHYLEARSLSNERDYDRDRYVDEYRNDYCEGVVPRHYHRD 92

QY 57 IESGYRIHCKSSVRSRSPKRNHCHSCSHQSRKSHRRKRSRSIEDDEGHLCQSG 116
DB 93 IESGYRIHCKSSVRSRSPKRNHCHSCSHQSRKSHRRKRSRSIEDDEGHLCQSG 152

QY 117 DVLARARYEIVDTLGGAGFKVVECDHGMGDMHVAVKIVKNGRYREARSEIQVLEHLN 176
DB 153 DVLARARYEIVDTLGGAGFKVVECDHGMGDMHVAVKIVKNGRYREARSEIQVLEHLN 212

QY 177 STDPSNPRCVQMLEFHHGHCIVFELLGLSTYDFIKENSFLPFQIDHROMAYQICQ 236
DB 213 STDPSNPRCVQMLEFHHGHCIVFELLGLSTYDFIKENSFLPFQIDHROMAYQICQ 272

QY 237 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDERTLKNNTDIKVDFGSAFYDDE 296
DB 273 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDERTLKNNTDIKVDFGSAFYDDE 332

QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 356
DB 333 HHSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 392

QY 357 LGPIPOHMIQTRKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCHDBEHEKLFDLV 416
DB 393 LGPIPOHMIQTRKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCHDBEHEKLFDLV 452

QY 417 RRMLEYDPTQRTITLDEALQHPFFDLKKK 445
DB 453 RRMLEYDPTQRTITLDEALQHPFFDLKKK 481

RESULT 6
US-09-810-671-4
; Sequence 4, Application US/09810671
; Publication No. US20020076783A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Human
US-09-810-671-4

Query Match      95.9%; Score 2312; DB 9; Length 427;
Best Local Similarity 100.0%; Pred. No. 5.8e-170;
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 HYLEARSLSNERDYDRDRYVDEYRNDYCEGVVPRHYHRDIESGYRIHCKSSVRSRSPK 78
DB 1 HYLEARSLSNERDYDRDRYVDEYRNDYCEGVVPRHYHRDIESGYRIHCKSSVRSRSPK 60

QY 79 RKRHRHCHSCSHQSRKSHRRKRSRSIEDDEGHLCQSGDVLARARYEIVDTLGGAGFKV 138
```

```

Db      61  RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRLARYEIVDTLGEAGFKV 120
Qy      139  ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 198
Db      121  ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 180
Qy      199  VCIVFELLGLSTYDFIKENSLPFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 258
Db      181  VCIVFELLGLSTYDFIKENSLPFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 240
Qy      259  VKSDYVVKYNSKMKRDEKTLKNTDIKVDFGSGATYDDDEHSHSTLVSTRHYRAPEVILGW 318
Db      241  VKSDYVVKYNSKMKRDEKTLKNTDIKVDFGSGATYDDDEHSHSTLVSTRHYRAPEVILGW 300
Qy      319  SOPCDVWSIGCILLEYLIGFTVQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 378
Db      301  SOPCDVWSIGCILLEYLIGFTVQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 360
Qy      379  LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 438
Db      361  LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 420
Qy      439  FDLKKK 445
Db      421  FDLKKK 427

RESULT 7
US-10-109-854-4
; Sequence 4, Application US/10109854
; Publication No. US200201195481
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-109-854-4

```

```

Query Match      95.9%; Score 2312; DB 13; Length 427;
Best Local Similarity 100.0%; Pred. No. 5.8e-170;
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19  HYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRDIESGYRIHCKSSVRSRSPK 78
Db      1  HYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRDIESGYRIHCKSSVRSRSPK 60
Qy      79  RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRLARYEIVDTLGEAGFKV 138
Db      61  RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRLARYEIVDTLGEAGFKV 120
Qy      139  ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 198
Db      121  ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 180
Qy      199  VCIVFELLGLSTYDFIKENSLPFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 258
Db      181  VCIVFELLGLSTYDFIKENSLPFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 240
Qy      259  VKSDYVVKYNSKMKRDEKTLKNTDIKVDFGSGATYDDDEHSHSTLVSTRHYRAPEVILGW 318
Db      241  VKSDYVVKYNSKMKRDEKTLKNTDIKVDFGSGATYDDDEHSHSTLVSTRHYRAPEVILGW 300
Qy      319  SOPCDVWSIGCILLEYLIGFTVQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 378
Db      301  SOPCDVWSIGCILLEYLIGFTVQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 360
Qy      379  LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 438
Db      361  LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 420

```

```

Db      241  VKSDYVVKYNSKMKRDEKTLKNTDIKVDFGSGATYDDDEHSHSTLVSTRHYRAPEVILGW 300
Qy      319  SOPCDVWSIGCILLEYLIGFTVQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 378
Db      301  SOPCDVWSIGCILLEYLIGFTVQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 360
Qy      379  LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 438
Db      361  LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 420
Qy      439  FDLKKK 445
Db      421  FDLKKK 427

RESULT 8
US-10-339-656-4
; Sequence 4, Application US/10339656
; Publication No. US20030134319A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/339,656
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-339-656-4

```

```

Query Match      95.9%; Score 2312; DB 14; Length 427;
Best Local Similarity 100.0%; Pred. No. 5.8e-170;
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19  HYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRDIESGYRIHCKSSVRSRSPK 78
Db      1  HYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRDIESGYRIHCKSSVRSRSPK 60
Qy      79  RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRLARYEIVDTLGEAGFKV 138
Db      61  RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRLARYEIVDTLGEAGFKV 120
Qy      139  ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 198
Db      121  ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 180
Qy      199  VCIVFELLGLSTYDFIKENSLPFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 258
Db      181  VCIVFELLGLSTYDFIKENSLPFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 240
Qy      259  VKSDYVVKYNSKMKRDEKTLKNTDIKVDFGSGATYDDDEHSHSTLVSTRHYRAPEVILGW 318
Db      241  VKSDYVVKYNSKMKRDEKTLKNTDIKVDFGSGATYDDDEHSHSTLVSTRHYRAPEVILGW 300
Qy      319  SOPCDVWSIGCILLEYLIGFTVQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 378
Db      301  SOPCDVWSIGCILLEYLIGFTVQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 360
Qy      379  LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 438
Db      361  LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 420

```



QY 439 FDLKKK 445  
Db 421 FDLKKK 427

## RESULT 9

US-10-801-671-4  
; Sequence 4, Application US/10801671  
; Publication No. US20040152123A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEROF  
; FILE REFERENCE: CLO00758DIV-111  
; CURRENT APPLICATION NUMBER: US/10/801,671  
; PRIOR FILING DATE: 2004-03-17  
; PRIOR APPLICATION NUMBER: 60/227,470  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 09/810,671  
; PRIOR FILING DATE: 2001-03-19  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 427  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-801-671-4

Query Match 95.9%; Score 2312; DB 16; Length 427;  
Best Local Similarity 100.0%; Pred. No. 5.8e-170;  
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 19 HYLEARSINERDYDRYVDEYNDYCEGYVPRHYRDIESGYRTHCSKSSVRSRSPK 78  
Db 1 HYLEARSINERDYDRYVDEYNDYCEGYVPRHYRDIESGYRTHCSKSSVRSRSPK 60  
QY 79 RKNRHCSSHQSRKSHRRKRSRSIEDDEGHLCQSGDVLRLRYEIVDTLGEAGFGKV 138  
Db 61 RKNRHCSSHQSRKSHRRKRSRSIEDDEGHLCQSGDVLRLRYEIVDTLGEAGFGKV 120  
QY 139 ECIDHGMGMVAVKIVNKGRYREARSEIQVLEHNLSTDPNSVFCVQMLEWFDHGH 198  
Db 121 ECIDHGMGMVAVKIVNKGRYREARSEIQVLEHNLSTDPNSVFCVQMLEWFDHGH 180  
QY 199 VCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYOICQINFLHNKLTHTDLKPENILF 258  
Db 181 VCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYOICQINFLHNKLTHTDLKPENILF 240  
QY 259 VKSDYVVKYNSKMKRDERTLKNITDVKVDYDFGSATYDDDEHSTLSTVSTRHYRAPEVILALGW 318  
Db 241 VKSDYVVKYNSKMKRDERTLKNITDVKVDYDFGSATYDDDEHSTLSTVSTRHYRAPEVILALGW 300  
QY 319 SQPCDWSIGCILIEYILGFTVFQTHDSKEHLAMMERILGPIQMIQTRKRYFHHNQ 378  
Db 301 SQPCDWSIGCILIEYILGFTVFQTHDSKEHLAMMERILGPIQMIQTRKRYFHHNQ 360  
QY 379 LDWDEHSSAGRYVRRCKPLKEFMLCHDEEHEKLPDLVRRMLEYDPTORITLDEALQHPF 438  
Db 361 LDWDEHSSAGRYVRRCKPLKEFMLCHDEEHEKLPDLVRRMLEYDPTORITLDEALQHPF 420  
QY 439 FDLKKK 445  
Db 421 FDLKKK 427

## RESULT 10

US-09-905-999-25  
; Sequence 25, Application US/09905999  
; Patent No. US2002010671A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, Axel  
; APPLICANT: NAYLOR, Oliver

; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS  
; FILE REFERENCE: 038602/0431  
; CURRENT APPLICATION NUMBER: US/09/905,999  
; CURRENT FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: 09/127,248  
; PRIOR FILING DATE: 1999-07-31  
; PRIOR APPLICATION NUMBER: PCI/IB97/00946  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 08/877,150  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 60/034,286  
; PRIOR FILING DATE: 1996-12-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 25  
; LENGTH: 481  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-905-999-25  
Query Match 93.8%; Score 2261.5; DB 9; Length 481;  
Best Local Similarity 93.5%; Pred. No. 5.2e-166;  
Matches 420; Conservative 8; Mismatches 10; Indels 11; Gaps 2;  
QY 8 SH-SVEEDTH-----PSHYLEARSINERDYDRYVDEYNDYCEGYVPRHYRD 56  
Db 33 SHSSTQENRCHCKPHQFKDSDCHYLEARCLNERDYDRYVDEYNDYCEGYVPRHYRD 92  
QY 57 IESGYRTHCSKSSVRSRSPKRNHRCSHQSRKSHRRKRSRSIEDDEGHLCQSG 116  
Db 93 VESTYRTHCSKSSVRSRSPKRNHRCSHQSRKSHRRKRSRSIEDDEGHLCQSG 152  
QY 117 DVLRLRYEIVDTLGEAGFGKVVECTIDHGMGMVAVKIVNKGRYREARSEIQVLEHNL 176  
Db 153 DVLRLRYEIVDTLGEAGFGKVVECTIDHGMGMVAVKIVNKGRYREARSEIQVLEHNL 212  
QY 177 STDNSVFCVQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYOICQ 236  
Db 213 STDNSVFCVQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYOICQ 272  
QY 237 SINFHNKLTHTDLKPENILFVKSDYVVKYNSKMKRDERTLKNITDVKVDYDFGSATYDD 296  
Db 273 SINFHNKLTHTDLKPENILFVKSDYVVKYNSKMKRDERTLKNITDVKVDYDFGSATYDD 332  
QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDWSIGCILIEYILGFTVFQTHDSKEHLAMMERI 356  
Db 333 HHSTLVSTRHYRAPEVILALGWSQPCDWSIGCILIEYILGFTVFQTHDSKEHLAMMERI 392  
QY 357 LGPIQMIQTRKRYFHHNQLDWDEHSSAGRYVRRCKPLKEFMLCHDEEHEKLPDLV 416  
Db 393 LGPIQMIQTRKRYFHHNQLDWDEHSSAGRYVRRCKPLKEFMLCHDEEHEKLPDLV 452  
QY 417 RMLEYDPTORITLDEALQHPFOLLKKK 445  
Db 453 RMLEYDPTORITLDEALQHPFOLLKKK 481  
RESULT 11  
US-10-267-502-355  
; Sequence 355, Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; APPLICANT: Galant, Ron  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 355  
; LENGTH: 481  
; TYPE: PRT



```

Db      361 HDRLDWDEHSSAGRYVSRAKPLKEFMLSQDVEHERLFDLIQKMLEYDPAKRITLTREALK 420
QY      436 HPFFDLLKK 444
         |||||
Db      421 HPFFDLLKK 429

RESULT 14
US-10-109-854-5
; Sequence 5, Application US/10109854
; Publication No. US20020119548A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-109-854-5

Query Match      78.3%; Score 1887; DB 13; Length 429;
Best Local Similarity 82.1%; Pred. No. 3.4e-137;
Matches 352; Conservative 30; Mismatches 45; Indels 2; Gaps 2;

QY      18 SHYLEASLNEDRYDRRYDEYRNDYCEGVVPRHYRDIESGVRHCSSKSSVSRSSP 77
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 SHYLESRSINEKDYHSRYIDEYRNDYTQCGEPGHRQRDHESRYQNHSSKSSGSRSSY 60
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      78 KRK-RNRHCSSH-QSRKSKSHRRKESRSIEDDEEGHLICQSGDVLARAYEIVDTLGEAGF 135
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 KSKHRIHSTSHRRSGKSHRRKTRSVEDDEEGHLICQSGDVLARSAYEIVDTLGEAGF 120
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      136 KVVCEIDHGMGDMHVAVKIKVNGVRYREAAARSEIQVLEHLNSTDPSNVSFRVQMLEWFDH 195
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      121 KVVCEIDHKAGGRHVAVKIKVNDRYCEAAARSEIQVLEHLNTDPSNSTRFCVQMLEWFEH 180
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      196 HGHCIVFELLGLSTYDFIKENSPLPQIDHIROMAYQICOSINFLHNKLTHTDLPEN 255
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      181 HGHCIVFELLGLSTYDFIKENGLPFLPDHIRKWAYQICKSVNFLSHNKLTHTDLPEN 240
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      256 ILFVKSDDYVVKYNSKMKRDERTLKNTDIKVDFGSATYDDEHHS TLVSTRHYRAPEVILA 315
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      241 ILFVQSDYTEAYNPKIKRDERTLNPDIKVDFGSATYDDEHHS TLVSTRHYRAPEVILA 300
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      316 LGWSQPCDWSIGCILLEYLGLTFVFTDHSKEHLAMMERILGPIQHMIOKTRKRYFH 375
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      301 LGWSQPCDWSIGCILLEYLGLTFVFTDHSKEHLAMMERILGPLPKHMIOKTRKRYFH 360
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      376 HNQLDWDEHSSAGRYVRRCKPLKEFMLCHDDEHEKLPDLVRRMLEYDPTQRIITLDEALQ 435
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      361 HDRLDWDEHSSAGRYVSRAKPLKEFMLSQDVEHERLFDLIQKMLEYDPAKRITLTREALK 420
         |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      436 HPFFDLLKK 444
         |||||
Db      421 HPFFDLLKK 429

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 23:51:06 ; Search time 170 Seconds

(without alignments)  
1340.443 Million cell updates/sec

Title: US-10-801-671-2

Perfect score: 2410

Sequence: 1 MCIPLEASHSVEDTHPSHY.....QRITLDELQHPFFDLKKK 445

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match % | ID           | Description         |
|------------|--------|---------------|--------------|---------------------|
| 1          | 2315.5 | 96.1          | 1 CLK4_HUMAN | Q9haz1 homo sapien  |
| 2          | 2261.5 | 93.8          | 1 CLK4_MOUSE | O35493 mus musculus |
| 3          | 1893   | 78.5          | 2 Q8N5V8     | Q8N5V8 homo sapien  |
| 4          | 1887   | 78.3          | 1 CLK1_HUMAN | P49759 homo sapien  |
| 5          | 1879   | 78.0          | 2 Q9BGQ1     | Q9bgq1 macaca fasc  |
| 6          | 1832   | 76.0          | 1 CLK1_MOUSE | P22518 mus musculus |
| 7          | 1803   | 74.8          | 2 Q9NRK6     | Q9nr16 homo sapien  |
| 8          | 1608   | 66.7          | 2 Q6AYK7     | Q6ayk7 rattus norv  |
| 9          | 1437   | 59.6          | 2 Q91YR2     | Q91yr2 mus musculus |
| 10         | 1427.5 | 59.2          | 1 CLK2_MOUSE | O35491 mus musculus |
| 11         | 1417   | 58.8          | 1 CLK2_HUMAN | P49760 homo sapien  |
| 12         | 1398   | 58.0          | 2 Q7ZV59     | Q7zy59 xenopus lae  |
| 13         | 1321   | 54.8          | 1 CLK3_HUMAN | P49761 homo sapien  |
| 14         | 1317   | 54.6          | 1 CLK3_RAT   | Q63117 rattus norv  |
| 15         | 1316   | 54.6          | 2 Q6IRK2     | Q6irk2 rattus norv  |
| 16         | 1310   | 54.4          | 1 CLK3_MOUSE | O35492 mus musculus |
| 17         | 1305   | 54.1          | 2 Q8C1V1     | Q8civ1 mus musculus |
| 18         | 1282   | 52.0          | 2 Q9BRG8     | Q9brg8 homo sapien  |
| 19         | 1180   | 49.0          | 3 D0A_DROME  | P49762 drosophila   |
| 20         | 1180   | 49.0          | 2 Q6GB73     | Q6gb73 drosophila   |
| 21         | 1122   | 46.6          | 3 Q7PN19     | Q7pn19 anopheles g  |
| 22         | 1119   | 46.4          | 3 Q7Q4W1     | Q7q4w1 anopheles g  |
| 23         | 1011   | 42.0          | 2 Q7PGH8     | Q7pgh8 anopheles g  |
| 24         | 969.5  | 40.2          | 2 Q1917      | Q1917 caenorhabdi   |
| 25         | 963    | 40.0          | 2 Q8MLY2     | Q8mly2 caenorhabdi  |
| 26         | 884.5  | 36.7          | 2 Q7PH65     | Q7ph65 anopheles g  |
| 27         | 848.5  | 35.2          | 1 AFC2_ARATH | P51567 arabidopsis  |
| 28         | 839    | 34.8          | 2 Q49967     | Q49967 nicotiana t  |
| 29         | 830    | 34.4          | 1 AFCL_ARATH | P51566 arabidopsis  |
| 30         | 802    | 33.3          | 2 Q943M7     | Q943m7 oryza sativ  |
| 31         | 790    | 32.8          | 1 AFC3_ARATH | P51568 arabidopsis  |

## ALIGNMENTS

| RESULT 1 | CLK4_HUMAN | STANDARD; | PRT; | 481 AA.      |
|----------|------------|-----------|------|--------------|
| 32       | 789        | 32.7      | 437  | 2 Q9M598     |
| 33       | 767        | 31.8      | 575  | 1 LKH1_SCHPO |
| 34       | 718.5      | 29.8      | 737  | 1 KNS1_YEAST |
| 35       | 717.5      | 29.8      | 658  | 2 Q6CD14     |
| 36       | 696        | 28.9      | 725  | 2 Q6FVQ2     |
| 37       | 690        | 28.6      | 640  | 2 Q6BU28     |
| 38       | 690        | 28.6      | 699  | 2 Q6CU13     |
| 39       | 673.5      | 27.9      | 772  | 2 Q7RZY1     |
| 40       | 672.5      | 27.9      | 673  | 2 Q7SAG1     |
| 41       | 648.5      | 26.9      | 881  | 2 Q8IL19     |
| 42       | 648.5      | 26.9      | 881  | 2 Q9BKN8     |
| 43       | 631.5      | 26.2      | 733  | 2 Q7B9I8     |
| 44       | 598        | 24.8      | 1300 | 2 Q8XOV5     |
| 45       | 550.5      | 22.8      | 538  | 2 Q8BM34     |

OS Homo sapiens (Human);  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AUTOPHOSPHORYLATION, AND MUTAGENESIS OF LYS-189.  
RC TISSUE=Kidney;  
RX MEDLINE=21100912; PubMed=11170754; DOI=10.1106/geno.2000.6447;  
RA Schultz J., Jones T., Bork P., Sheer D., Blencke S., Steyrer S.,  
RA Wellbrock U., Bevec D., Ullrich A., Wallasch C.;  
RT "Molecular characterization of a cDNA encoding functional human CLK4 kinase and localization to chromosome 5q35.";  
RL Genomics 71:368-370(2001).  
RN [2]  
RN ERRATUM.  
RP Schultz J., Jones T., Bork P., Sheer D., Blencke S., Steyrer S.,  
RA Wellbrock U., Bevec D., Ullrich A., Wallasch C.;  
RL Genomics 74:251-251(2001).  
RN [3]  
RN INTERACTION WITH UBL5.  
RP MEDLINE=22709234; PubMed=12824502; DOI=10.1110/ps.0382803;  
RA McNally T., Huang Q., Janis R.S., Liu Z., Olejniczak E.T.,  
RA Reilly R.M.;  
RT "Structural analysis of UBL5, a novel ubiquitin-like modifier.";  
RL Protein Sci. 12:1562-1566(2003).  
RN [4]  
RN INTERACTION WITH UBL5.  
RP MEDLINE=22592248; PubMed=12705895; DOI=10.1016/S0006-291X(03)00549-7;  
RA Kantham L., Kerr-Bayles L., Godde N., Quick M., Webb R.,  
RA Sunderland T., Bond J., Walder K., Augert G., Collier G.;  
RT "Beacon interacts with cdc2/cdc28-like kinases.";  
RL Biochem. Biophys. Res. Commun. 304:125-129(2003).  
CC -!- FUNCTION: Phosphorylates serine- and arginine-rich (SR) proteins of the spliceosomal complex may be a constituent of a network of regulatory mechanisms that enable SR proteins to control RNA splicing. Phosphorylates serines, threonines and tyrosines.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
CC -!- SUBUNIT: Interacts with UBL5.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -!- TISSUE SPECIFICITY: Expressed in liver, kidney, heart, muscle and

Q9m598 mesembryant  
Q10156 schizosacch  
P32350 saccharomyc  
Q6cd14 yarrowia li  
Q6fvq2 candida gla  
Q6bu28 debaryomyc  
Q6cjl3 kluyveromyc  
Q7rzy1 neurospora  
Q7sag1 ashbya goss  
Q8il19 plasmodium  
Q9bkn8 plasmodium  
Q7b9i8 plasmodium  
Q8xov5 neurospora  
Q8bm34 m mus muscu



RA Butterfield Y.S.N., Krzywinaki M.J., Skalska U., Smailus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 1-219 FROM N.A.  
 RX MEDLINE=97480726; PubMed=9339371; DOI=10.1006/geno.1997.4931;  
 RA Watkins-Chow D.E., Douglas K.R., Buckwalter M.S., Probst F.J.,  
 RA Camper S.A.;  
 RT "Construction of a 3-Mb contig and partial transcript map of the  
 RT central region of mouse chromosome 11.";  
 RL Genomics 45:147-157(1997).  
 RN [5]  
 RP FUNCTION.  
 RX MEDLINE=21354768; PubMed=11461155; DOI=10.1006/mcne.2001.1000;  
 RA Hartmann A.M., Rujescu D., Giannakouros T., Nikolaki E., Goedert M.,  
 RA Mandelkow E.-M., Gao Q.S., Andreadis A., Stamm S.;  
 RT "Regulation of alternative splicing of human tau exon 10 by  
 RT phosphorylation of splicing factors.";  
 RL Mol. Cell. Neurosci. 18:80-90(2001).  
 RN [6]  
 RP TISSUE SPECIFICITY, AND MUTAGENESIS OF LYS-189.  
 RX MEDLINE=22313485; PubMed=12169693; DOI=10.1074/jbc.M206504200;  
 RA Katsu R., Onogi H., Wada K., Kawaguchi Y., Hagiwara M.;  
 RT "Novel SR-rich-related protein Clasp specifically interacts with  
 RT inactivated Clk4 and induces the exon EB inclusion of Clk.";  
 RL J. Biol. Chem. 277:44220-44228(2002).  
 CC -!- FUNCTION: Phosphorylates serine- and arginine-rich (SR) proteins  
 CC of the spliceosomal complex may be a constituent of a network of  
 CC regulatory mechanisms that enable SR proteins to control RNA  
 CC splicing. Phosphorylates serines, threonines and tyrosines.  
 CC Required for the regulation of alternative splicing of MAPT/TAU.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBUNIT: Interacts with UBL5 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=O35493-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=O35493-2; Sequence=VSP\_008205;  
 CC -!- TISSUE SPECIFICITY: Expressed in the hippocampus, the cerebellum  
 CC and the olfactory bulb.  
 CC -!- PTM: Autophosphorylates on all three types of residues.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Lammer  
 CC subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF033566; AB87510.1; -;  
 CC EMBL; AK013974; BAC23420.1; -;  
 CC EMBL; BC002220; AAH02220.1; -;  
 CC EMBL; BC012675; AAH12675.1; -;  
 CC EMBL; U94846; AAB62179.1; -;  
 CC HSSP; Q00534; 1B18.  
 CC MGD; MGI:1098551; Clk4.  
 CC InterPro; IPR011009; Kinase like.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 CC InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC Pfam; PF00069; Pkinase; 1.  
 CC ProDom; PD000001; Prot\_kinase; 1.  
 CC SMART; SM00220; S\_TK; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Alternative splicing; ATP-binding; Nuclear protein; Phosphorylation;  
 KW Serine/threonine-protein kinase; Transferase; Tyrosine-protein kinase.  
 FT DOMAIN 159 475 Protein kinase.  
 FT NP BIND 165 173 ATP (By similarity).  
 FT BINDING 189 189 ATP (By similarity).  
 FT ACT\_SITE 286 286 Proton acceptor (By similarity).  
 FT VARSPLIC 1 180 MISSING (in isoform 2).  
 FT /FTID=VSP\_008205.  
 FT MUTAGEN 189 189 K->R: Loss of function.  
 FT CONFLICT 58 58 E -> Q (in Ref. 4).  
 FT CONFLICT 141 141 D -> N (in Ref. 4).  
 FT CONFLICT 344 344 R -> K (in Ref. 2).  
 SQ SEQUENCE 481 AA; 57344 MW; F2C56965900C12AA CRC64;  
 Query Match 93.8%; Score 2261.5; DB 1; Length 481;  
 Best Local Similarity 93.5%; Pred No. 5.1e-132;  
 Matches 420; Conservative 8; Mismatches 10; Indels 11; Gaps 2;  
 QY 8 SH-SVEEDTH-----PSHYLEARSLSNERDYRDYRVDYDEYRNDYCEGYVPRHYRD 56  
 DB 33 SHSSTQENRHCKPHQKXDSCHYLEARCLNERDYRDYRVDYDEYRNDYCEGYVPRHYRD 92  
 QY 57 IESGRIHCSKSVSRSSSPKRNHCHSSHQSSKSHRRKRSRSIEDDEGHLCQSG 116  
 DB 93 VESTYRIHCSKSVSRSSSPKRNPCASHQSHSKSHRRKRSRSIEDDEGHLCQSG 152  
 QY 117 DVLRYRIYDVLGEGAGKGVVECDHGMGMHVAVKIVKNVGRYREAAARSETQVLEHLN 176  
 DB 153 DVLRYRIYDVLGEGAGKGVVECDHGMGMHVAVKIVKNVGRYREAAARSETQVLEHLN 212  
 QY 177 STDPSNFRVCQMLEWFHGHGHCIVFELLGLSTYDFIKENSLFPPQIDHIRQMAVQICQ 236  
 DB 213 STDPSNFRVCQMLEWFHGHGHCIVFELLGLSTYDFIKENSLFPPQIDHIRQMAVQICQ 272  
 QY 237 SINFLHNKLTHTDLKPENILFKVSDYVVKYNSKMKRDERTKNTDIKVDVFGSATYDDE 296  
 DB 273 SINFLHNKLTHTDLKPENILFKVSDYVVKYNSKMKRDERTKNTDIKVDVFGSATYDDE 332  
 QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYVLGFTVQTHDSKHLAMMERI 356  
 DB 333 HHSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYVLGFTVQTHDSKHLAMMERI 392  
 QY 357 LGPIQPMIQTCKRKYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLV 416  
 DB 393 LGPIQPMIQTCKRKYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLV 452  
 QY 417 RMLSYDPTQRTILDEALQHPFDLLKKX 445  
 DB 453 RMLSYDPTQRTILDEALQHPFDLLKKX 481  
 RESULT 3  
 QN5V8  
 ID QN5V8 PRELIMINARY; PRT; 484 AA.  
 AC QN5V8;  
 DT 01-OCT-2002 (TREMELrel. 22, Created)  
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
 DE CDC-like kinase 1.  
 GN Name=CLK1;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,





```

DR PTR; S53641; S53641.
DR HSP; Q00534; 1B18.
DR Genew; HGNC:2068; CLK1.
DR H-InvDB; HIX0002728; -.
DR MIM; 601951; -.
DR GO; Q0004715; F:non-membrane spanning protein tyrosine kinase; TAS.
DR GO; Q0004674; P:protein serine/threonine kinase activity; TAS.
DR GO; Q0000074; P:cell proliferation; TAS.
DR GO; Q0000074; P:regulation of cell cycle; TAS.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr pkin_AS.
DR InterPro; IPR002290; Ser Thr pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW Alternative splicing; ATP-binding; Nuclear protein; Phosphorylation;
Serine/threonine-protein kinase; Transferase; Tyrosine-protein kinase.
FT DOMAIN 161 477
FT NP_BIND 167 175 ATP (By similarity).
FT BINDING 191 191 ATP (By similarity).
FT ACT_SITE 288 288 Proton acceptor (By similarity).
FT VARSPPLIC 131 136 KSHRRK -> MKLLIL (in isoform Short).
FT VARSPPLIC 137 484 Missing (in isoform Short).
FT VARSPPLIC 137 484 /FTId=VSP_004852.
FT SEQUENCE 484 AA; 57205 MW; 304958486AD0A6B CRC64;
Query Match 78.3%; Score 1887; DB 1; Length 484;
Best Local Similarity 82.1%; Pred. No. 7.4e-109;
Matches 352; Conservative 30; Mismatches 45; Indels 2; Gaps 2;
QY 18 SHYLEARSLSNERDYRRDYVDEYRNDYCEGYVPRPHYRDISSYRTHCSKSSVRSRSP 77
DB 54 SHYLESRSINEXDYGRRYIDEYRNDYTOGCEPGRHQRDESHYQNHSSKSGRSY 113
QY 78 KRR-RNRHCSSH-QSRKSHRRKRSSIEDDEGHLCIQSGDVLRYEIVDTLGGAGF 135
DB 114 KSKRHHTSHSTRSHGSKSHRRKTRTSVEDDEGHLCIQSGDVLRYEIVDTLGGAGF 173
QY 136 KVECIDHGMGMHVAVKIVNRYREARSIOVLEHLNSTDPNSVFRVQMLEWFDH 195
DB 174 KVECIDHGRGRHVAVKIVNDRYCEARSIOVLEHLNSTDPNSTFRVQMLEWFEH 233
QY 196 HGVCIIVFELLGLSTYDFIKENSFLPQDHIROMAYQICQSNPLHNKLTHTDLKPEN 255
DB 234 HGHCIVFELLGLSTYDFIKENGFLPFLDHIRKMAQYQICSNVFLHNSNKLTHDLPEN 293
QY 256 ILFVKSQDVYVYKNSKMRDERTLKNITDIKVDFGSGATYDEHSTLVSTRHYRAPVILA 315
DB 294 ILFVQSDYTYEYNPKIKRDETLINFDIKVDFGSGATYDEHSTLVSTRHYRAPVILA 353
QY 316 LGWSQPCDVMSIGCILLIYVLTFTQTHDSKEHLAMMERILGPIQPMIQRKRKYFH 375
DB 354 LGWSQPCDVMSIGCILLIYVLTFTQTHDSKEHLAMMERILGPIQPMIQRKRKYFH 413
QY 376 HNQLDWEHSSAGRYVRRCKPKLFKMLCHDEBEKLFDLVRMLYDPTQRTILDEALQ 435
DB 414 HDRLDWEHSSAGRYVSRACKPKLFKMLSDQVEHERLFDLIQKMLEYDPAKRITLREALK 473
QY 436 HPFFDLKK 444
DB 474 HPFFDLKK 482
RESULT 5
Q9BGQ1 PRELIMINARY; PRT; 484 AA.
ID Q9BGQ1
AC Q9BGQ1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

---

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal lobe left;
RA Ceada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB056421; BAB33079.1; -.
DR HSP; Q00534; 1B18.
DR GO; Q0005524; P:ATP binding; IEA.
DR GO; Q0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; Q0016740; P:transferase activity; IEA.
DR GO; Q0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr pkinase.
DR InterPro; IPR008271; Ser Thr pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; UNKNOWN 1.
KW ATP-binding; Hypothetical protein; Kinase; Transferase.
SQ SEQUENCE 484 AA; 57318 MW; F5324D4007119878 CRC64;
Query Match 78.0%; Score 1879; DB 2; Length 484;
Best Local Similarity 79.3%; Pred. No. 2.3e-108;
Matches 356; Conservative 30; Mismatches 51; Indels 12; Gaps 4;
QY 8 SHSVEED-----TH-----PSHYLEARSLSNERDYRRDYVDEYRNDYCEGYVPRPHYRDI 57
DB 34 SHSSARENRCKYTHSKMCDSHYLESRCINEXDYGRRYIDEYRNDYNOCEPGRHHRDH 93
QY 58 ESGYRTHCSKSSVRSRSPKPK-RNRHCSSH-QSRKSHRRKRSSIEDDEGHLCIQS 115
DB 94 ESGYQNHSSKSGRSRSGSYKSKHRTHSTSHRRSHGSKSHRRKTRTSVEDDEGHLCIQS 153
QY 116 GDLVRLARYEIVDTLGGAGFQVVECIDHGMGMHVAVKIVNRYREARSIOVLEHL 175
DB 154 GDLVSARYEIVDALGEGAGFQVVECIDHKGAGHVAVKIVNDRYCEARSIOVLEHL 213
QY 176 NSTDPNSVFRVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQDHIROMAYQIC 235
DB 214 NTTDPNSTFRVQMLEWFEHGHGHCIVFELLGLSTYDFIKENGFLPFLDHIRKMAQYQIC 273
QY 236 QSNFLHNKLTHTDLKPENILFVKSQDVYVYKNSKMRDERTLKNITDIKVDFGSGATYDD 295
DB 274 KSNVFLHNSNKLTHDLPENILFVQSDYTYEYNPKIKRDETLINFDIKVDFGSGATYDD 333
QY 296 EHHSTLVSTRHYRAPVILALGWSQPCDVMSIGCILLIYVLTFTQTHDSKEHLAMMER 355
DB 334 EHHSTLVSTRHYRAPVILALGWSQPCDVMSIGCILLIYVLTFTQTHDSKEHLAMMER 393
QY 356 ILGPIQPMIQRKRKYFHNLQDWEHSSAGRYVRRCKPKLFKMLCHDEBEKLFDL 415
DB 394 ILGLPKPMIQRKRKYFHNDRLDWEHSSAGRYVRRCKPKLFKMLSDQVEHECLFDL 453
QY 416 VRMLYDPTQRTILDEALQHPFFDLKK 444
DB 454 IQKMLEYDPAKRITLKEALKHPFFDLKK 482
RESULT 6
CLK1_MOUSE STANDARD; PRT; 483 AA.
ID CLK1_MOUSE

```









|    |     |   |     |
|----|-----|---|-----|
| Db | 290 | DLKPENILFVNSDYELTYNLEKKKRSVKSTAVRVVDFGSAFTDHEHSTIVSTRHYRA     | 349 |
| Qy | 310 | PEVILALGWQPCDQWSIGCILEYVIGFTVFOTHDSKEHLAMMERILGIPQHMIOKTR     | 369 |
| Db | 350 | PEVILELWGWQPCDQWSIGCILEYVVGFTVFOTHDRREHLAMMERILGIPGRMTIKTR    | 409 |
| Qy | 370 | KRYFYRHNQLDWDHSSAGRYVRRRCFKLFEKFMCLHDEEKEKFLDLVRRMLRYDPTQIRIT | 429 |
| Db | 410 | KQKVFYRGRGLDWDENTSAGRYVRENCPLRRYLYTSEAEHHQLFDLIESMLEVEPAKRLT  | 469 |
| Qy | 430 | LDLALQHPFFDLIK  | 443 |
| Db | 470 | LGEALQHPFFARLR  | 483 |

RESULT 12

Q7ZY59 PRELIMINARY; PRT; 491 AA.

AC Q7ZY59

AT Q7ZY59

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

OS Clk2-prov protein.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodinae; Xenopus.

OC NCBI\_TaxID=8355;

RN [1]

RN SEQUENCE FROM N.A.

RP TISSUE=Embryo;

RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,

RA Krzywiński M.I., Skalska U., Smalius D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RN SEQUENCE FROM N.A.

RP TISSUE=Embryo;

RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;

RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

RA Richardson P.;

RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RT initiative."

RT Dev. Dyn. 225:384-391 (2002).

RN [3]

RN SEQUENCE FROM N.A.

RP TISSUE=Embryo;

RC Klein S., Strausberg R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

RL EMBL; BC043963; AAH43963.1; --

DR HSP; Q03456; 1HOW.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.

DR GO; GO:0016740; P:transferase activity; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPRO11009; Kinase like.

DR InterPro: IPR000719; Prot kinase.  
DR InterPro: IPR002290; Ser thr kinase.  
DR InterPro: IPR008271; Ser thr pkin\_AS.  
DR Pfam: PF00069; Pkinase; 1.  
DR ProDom: PD000001; Prot kinase; 1.  
DR SMART: SM00220; S\_TKC\_1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW ATP-binding; Kinase; Transferase.  
SQ SEQUENCE 491 AA; 56994 MW; 20870CC5962303DF CRC64;  
  
Query Match 58.0%; Score 1398; DB 2; Length 491;  
Best Local Similarity 59.5%; Pred. No. 1.3e-78;  
Matches 262; Conservative 68; Mismatches 94; Indels 16; Gaps 6;  
  
QY 6 EASHSVVEEDTHPSHYLEARNLNERDYRDYVDEV-RNDYCEGVPRHYHDIESGYRTH 64  
DB 50 ERSRSVEE-----RSSDRRAY-DRKYCDSYRNDYSDRDRGVYYETDYY--YK-H 95  
QY 65 CSKSSVRSRRSPKRNKRNCHSHQSRKSHR-RKRSRSSTEDDEEGLHCQSGDVLARY 123  
DB 96 SRDSDSYRSTRKQKRNRRTRSYQSQRSSRSQSSREKSVEDDVEGLHYHSGDWLQERY 155  
QY 124 EIVDTLGEAGKGVCEIDHGMGMHVAIVKNGRYREARSEIQVLEHLNSTDPNSV 183  
DB 156 EIVSTLGECTGFRVQCKDHRGGRSVALKIKNVEKYKEARLEINVLEKINEKDPENK 215  
QY 184 FRCVQMLEWFDHGHVCIVPELLGLSTYDFIKENSLPQIDHIROMAVQICOSINFLHH 243  
DB 216 HLCVQMFQWFDVHGHMCISPELLGLSTDFLKNNYFPYPIQVRRHVALQCOAMKFLHD 275  
QY 244 NKLTHTDLPENILFVKSDYVVKYNSKMKRDERTLKNTOIKVDFGSGATYDDEHSTLVS 303  
DB 276 NKLTHTDLPENILFVSSDYELRYNNEKKRDERCVKSTDIRVDFGSAFDEHHSTIVS 335  
QY 304 TRHYRAPEVILALGWSQPCDWSIGCILIEYLGLFTVFOTHDSKEHLAMMERILGIPDH 363  
DB 336 TRHYRAPEVLLLELGNQPCDWSVGCIIPEIYVVGFTFQTHDNREHLAMMERILGIPSR 395  
QY 364 MIQKTRKRYVHNQLDWDEHSAGRYVRRCKPLKEFMLCHDEEHEKLFDLVRRMLEYD 423  
DB 396 MIKTRKQKIFYHGRUDWDNDTSAGRYVRENCKPLRYMMETEEHHQFNLEIGLLGYE 455  
QY 424 PTQRTIDLEALQHPFFDLK 443  
DB 456 PSKEWTLAALAKHPFFNPLK 475  
[1]  
  
RESULT 13  
CLK3 HUMAN STANDARD; PRT; 490 AA.  
AC P49761; Q9BR53; Q9BU7;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Dual specificity protein kinase CLK3 (EC 2.7.1.37) (EC 2.7.1.112) (CDC like kinase 3).  
GN Name=CLK3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=95082033; PubMed=7990150;  
RA Hanes J.J., der Kammer H., Klaidiny J.J., Scheit K.H.;  
RT "Characterization by cDNA cloning of two new human protein kinases.  
RT Evidence by sequence comparison of a new family of mammalian protein kinases";  
RL J. Mol. Biol. 244:665-672(1994).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).

RC TISSUE=Cervix, and Placenta;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SPLICE ISIFORM(S) THAT ARE POTENTIAL NMD TARGET(S).  
RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;  
RA Hillman R.T., Green R.E., Brenner S.E.;  
RT "An unappreciated role for RNA surveillance.";  
RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).  
CC -!- FUNCTION: Phosphorylates serine- and arginine-rich (SR) proteins  
CC of the spliceosomal complex may be a constituent of a network of  
CC regulatory mechanisms that enable SR proteins to control RNA  
CC splicing. Phosphorylates serines, threonines and tyrosines (By  
CC similarity).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1; Synonyms=Long;  
CC IsoId=P49761-1; Sequence=Displayed;  
CC Name=2; Synonyms=Short;  
CC IsoId=P49761-2; Sequence=VSP\_004858, VSP\_004859;  
CC Notes=Lacks the kinase domain. May be produced at very low levels  
CC due to a premature stop codon in the mRNA, leading to  
CC nonsense-mediated mRNA decay;  
CC Name=3;  
CC IsoId=P49761-3; Sequence=VSP\_004860;  
CC -!- PTM: Autophosphorylates on all three types of residues (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Lammer  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L29217; AAA61484.1; -;  
DR EMBL; L29220; AAA61483.1; -;  
DR EMBL; BC002555; AAH02555.1; -;  
DR EMBL; BC006103; AAH06103.1; -;  
DR EMBL; BC019881; AAH19881.1; -;  
DR PIR; S53639; S53639.  
DR PIR; S53640; S53640.  
DR HSSP; Q00534; 1B18.  
DR Genew; HGNC:2071; CLK3.  
DR H-InvDB; HIX0018051; -;  
DR MIM; 602990; -;  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.  
DR



|           |          |  |
|-----------|----------|--|
| RESULT 14 | CLK3_RAT |  |
| ID        | _CLK3    |  |
| AC        | Q6311    |  |
| DT        | 15-D     |  |
| DT        | 15-D     |  |
| DT        | 05-J     |  |
| DE        | Dual     |  |
| DE        | like     |  |
| GN        | Name     |  |



```

Db 199 RLEINVLLKKIKKEDKENKFLCVLMSDFNFHGHMCIAPELLGKNTPEFLKENNFQYPLP 258
QY 226 HIROMAYQICQSFNFIHNNKLTHTDLKPNILFVKSDDYVVKNSKMKRDERLTAKNTDIKV 285
Db 259 HVHRMAYQLCHALRFLHFNQLTHTDLKPNILFVNSEFETLYNEHKSCEKSKVNTSIRV 318
QY 286 VFGSATYDDEHSTLSTVSTRHRAPEVILALGWSQPCDWSIGCILIEYLLGFTVPQTTHD 345
Db 319 ADFGSATFDHEHHTTTVATRHVRPPEVILELWQAQPCDWSIGCILIEYLLGFTVPQTTHD 378
QY 346 SKEHLAMMERILGPIPOHMIQKTRKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCH 405
Db 379 NREHLVMMKILGPIPSHMIHTRKQKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCH 438
QY 406 DEHEKLFDLVRRMLEYDPTQRTITLDEALQHPFF 439
Db 439 SLEHVQLFDLMRMLFDFPSQRTITLAEALLHPFF 472

RESULT 15
QID Q6IRK2 PRELIMINARY; PRT; 490 AA.
AC Q6IRK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Clk3 protein.
GN Name=Clk3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomii;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapletenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC070891; AAH70891.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr kinase.
DR InterPro; IPR008271; Ser Thr_pkin_A5.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.

```

```

DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; kinase; transferase.
SQ SEQUENCE 490 AA; 58485 MW; 6B10CA96A64AB19A CRC64;

Query Match 54.6%; Score 1316; DB 2; Length 490;
Best Local Similarity 55.1%; Pred. No. 1.6e-73;
Matches 250; Conservative 60; Mismatches 102; Indels 42; Gaps 6;

QY 4 PLEASHSVEDTHPSHYLEARSINERDYDRRYVVEYRNDYCEGYVP-----R 51
Db 43 PPRSRSRSHDRIP-----YQRYREHRSDTYR---CEERSPFGEDCYGSSRSR 90
QY 52 HYHRDIESG-YRI-----HCSKSSVRSRSPKRNHCHSSHQSRKSHRRKRSRSIED 105
Db 91 HRRSRERGPYRTRKHAHHCHK-----RRTSCSSASSRSQSSKRSRVED 138
QY 106 DEEGHLICQSGDVLARARYEIVDTLGEAGFKVVECIDHGMGMHVAVKIVKNGVYREA 165
Db 139 DKEGHLVCRIGDWLQERYEIVGNLGEETGKVECLDHARGKSQVALKIIRNVGKYREA 198
QY 166 RSEIQVLEHLNSTDPNSVRCVOMLEWFDDHGHVCIIVFELLGLSTYDFIKENSFLPFQID 225
Db 199 RLEINVLLKKIKKEDKENKFLCVLMSDFNFHGHMCIAPELLGKNTPEFLKENNFQYPLP 258
QY 226 HIROMAYQICQSFNFIHNNKLTHTDLKPNILFVKSDDYVVKNSKMKRDERLTAKNTDIKV 285
Db 259 HVHRMAYQLCHALRFLHFNQLTHTDLKPNILFVNSEFETLYNEHKSCEKSKVNTSIRV 318
QY 286 VFGSATYDDEHSTLSTVSTRHRAPEVILALGWSQPCDWSIGCILIEYLLGFTVPQTTHD 345
Db 319 ADFGSATFDHEHHTTTVATRHVRPPEVILELWQAQPCDWSIGCILIEYLLGFTVPQTTHD 378
QY 346 SKEHLAMMERILGPIPOHMIQKTRKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCH 405
Db 379 NREHLVMMKILGPIPSHMIHTRKQKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCH 438
QY 406 DEHEKLFDLVRRMLEYDPTQRTITLDEALQHPFF 439
Db 439 SLEHVQLFDLMRMLFDFPSQRTITLAEALLHPFF 472

```

Search completed: March 13, 2005, 00:08:00

Job time : 173 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 04:53:38 ; Search time 10125.3 Seconds  
(without alignments)  
11265.245 Million cell updates/sec

Title: US-10-801-671-1

Perfect score: 2354  
Sequence: 1 gccagctggggtacttta.....taaaaaaacaaaaaaac 2354

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.on.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID       | Description        |
|------------|--------|-------------|--------|-------------|--------------------|
| 1          | 2354   | 100.0       | 2354   | 6 AR232168  | AR232168 Sequence  |
| 2          | 2354   | 100.0       | 2354   | 6 AR406007  | AR406007 Sequence  |
| 3          | 2354   | 100.0       | 2354   | 6 AR534243  | AR534243 Sequence  |
| 4          | 2354   | 100.0       | 2354   | 6 AR392827  | AR392827 Sequence  |
| 5          | 2052.6 | 87.2        | 2446   | 6 AR077239  | AR077239 Sequence  |
| 6          | 2052.6 | 87.2        | 2446   | 6 BD107903  | BD107903 Protein k |
| 7          | 2048   | 87.0        | 2488   | 6 AX056410  | AX056410 Sequence  |
| 8          | 1632   | 69.3        | 1814   | 9 AF294429  | AF294429 Homo sapi |
| 9          | 1629.4 | 69.2        | 1792   | 6 AX961896  | AX961896 Sequence  |
| 10         | 1318.4 | 56.0        | 3492   | 9 BC063116  | BC063116 Homo sapi |
| 11         | 1284   | 54.5        | 1446   | 12 AY335726 | AY335726 Synthetic |
| 12         | 1277.6 | 54.3        | 1914   | 10 BC012675 | BC012675 Mus muscu |
| 13         | 1277.4 | 54.3        | 1881   | 6 AX961894  | AX961894 Sequence  |
| 14         | 1185   | 50.3        | 1222   | 6 BD260616  | BD260616 49 human  |
| 15         | 1178.2 | 50.1        | 1456   | 6 AR380154  | AR380154 Sequence  |
| 16         | 1118.8 | 47.5        | 1865   | 6 AX961898  | AX961898 Sequence  |
| 17         | 1097.4 | 46.6        | 1549   | 10 AF033566 | AF033566 Mus muscu |
| 18         | 1097   | 46.6        | 21234  | 6 AR232169  | AR232169 Sequence  |
| 19         | 1097   | 46.6        | 21234  | 6 AR406008  | AR406008 Sequence  |

|      |        |      |        |              |                    |
|------|--------|------|--------|--------------|--------------------|
| 20   | 1097   | 46.6 | 21234  | 6 AR534244   | AR534244 Sequence  |
| 21   | 1097   | 46.6 | 21234  | 6 AX392829   | AX392829 Sequence  |
| c 22 | 1097   | 46.6 | 159266 | 2 AC026097   | AC026097 Homo sapi |
| c 23 | 1097   | 46.6 | 170624 | 9 AC022096   | AC022096 Homo sapi |
| 24   | 1009.6 | 42.9 | 1445   | 10 BC002220  | BC002220 Mus muscu |
| 25   | 995    | 42.3 | 1150   | 6 CQ718110   | CQ718110 Sequence  |
| 26   | 956.8  | 40.6 | 1414   | 10 AY653308  | AY653308 Mus muscu |
| 27   | 910    | 38.7 | 1416   | 10 BC079006  | BC079006 Rattus no |
| 28   | 881.6  | 37.5 | 1745   | 5 BX931976   | BX931976 Gallus ga |
| 29   | 824.4  | 35.0 | 1864   | 9 AB056421   | AB056421 Macaca fa |
| 30   | 814.8  | 34.6 | 1757   | 6 CQ715616   | CQ715616 Sequence  |
| 31   | 813.2  | 34.5 | 1773   | 9 BC031549   | BC031549 Homo sapi |
| 32   | 811.6  | 34.5 | 1834   | 9 HUMCLK1B   | L29219 Homo sapien |
| 33   | 768.8  | 32.7 | 1705   | 10 NMCLMRNA  | X57186 Mouse clk m |
| 34   | 765.6  | 32.5 | 1740   | 10 MUYSTYKIN | M38381 Mouse serin |
| 35   | 692    | 29.4 | 1946   | 5 CR848121   | CR848121 Xenopus t |
| 36   | 677.8  | 28.8 | 1544   | 5 BX936001   | BX936001 Gallus ga |
| 37   | 669.4  | 28.4 | 1017   | 5 BX929601   | BX929601 Gallus ga |
| 38   | 644.4  | 27.4 | 1743   | 9 HUMCLK1A   | L29222 Homo sapien |
| 39   | 643.2  | 27.3 | 1750   | 9 HUMKINCD   | MS9287 Human prote |
| 40   | 640.6  | 27.2 | 2254   | 6 AR531674   | AR531674 Sequence  |
| c 41 | 639.4  | 27.2 | 1535   | 11 BV179987  | BV179987 sqm10751  |
| 42   | 630.6  | 26.8 | 1643   | 10 MUSWCLK   | L29221 Mus musculu |
| 43   | 628.2  | 26.7 | 4862   | 10 BC062887  | BC062887 Mus muscu |
| 44   | 617    | 26.2 | 3056   | 10 BC081942  | BC081942 Rattus no |
| 45   | 595.2  | 25.3 | 1634   | 5 BX929386   | BX929386 Gallus ga |

#### ALIGNMENTS

|            |   |                                    |         |     |        |                 |
|------------|---|------------------------------------|---------|-----|--------|-----------------|
| RESULT 1   | AR232168  | Sequence 1 from patent US 6455291. | 2354 bp | DNA | linear | PAT 20-DEC-2002 |
| LOCUS      | AR232168  | Sequence 1 from patent US 6455291. | 2354 bp | DNA | linear | PAT 20-DEC-2002 |
| DEFINITION | AR232168  | Sequence 1 from patent US 6455291. | 2354 bp | DNA | linear | PAT 20-DEC-2002 |
| ACCESSION  | AR232168  | Sequence 1 from patent US 6455291. | 2354 bp | DNA | linear | PAT 20-DEC-2002 |
| VERSION    | AR232168.1  | GI:27274056                        | 2354 bp | DNA | linear | PAT 20-DEC-2002 |
| KEYWORDS   | Unknown.  |                                    |         |     |        |                 |
| SOURCE     | Unknown.  |                                    |         |     |        |                 |
| ORGANISM   | Unknown.  |                                    |         |     |        |                 |
| REFERENCE  | 1 (bases 1 to 2354)   |                                    |         |     |        |                 |
| AUTHORS    | Yan, C., Ye, J., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.                                      |                                    |         |     |        |                 |
| TITLE      | Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof |                                    |         |     |        |                 |
| JOURNAL    | Patent: US 6455291-A 1 24-SEP-2002;   |                                    |         |     |        |                 |
| FEATURES   | Location/Qualifiers   |                                    |         |     |        |                 |
| source     | 1..2354   |                                    |         |     |        |                 |
| ORIGIN     | /organism="unknown"   |                                    |         |     |        |                 |
|            | /mol_type="genomic DNA"   |                                    |         |     |        |                 |

|                       |        |  |       |            |    |        |       |
|-----------------------|--------|--|-------|------------|----|--------|-------|
| Query Match           | 100.0% | Score  | 2354; | DB         | 6; | Length | 2354; |
| Best Local Similarity | 100.0% | Pred. No.  | 0;    |            |    |        |       |
| Matches               | 2354;  | Conservative   | 0;    | Mismatches | 0; | Indels | 0;    |
| Gaps                  | 0;     |  |       |            |    |        |       |
| Qy                    | 1      | GCCAGCTGGGGTTACTTTAAAAACATCTCCATGTGCATCCCTCTTGAAGCTTCGCACT   | 60    |            |    |        |       |
| Db                    | 1      | GCCAGCTGGGGTTACTTTAAAAACATCTCCATGTGCATCCCTCTTGAAGCTTCGCACT   | 60    |            |    |        |       |
| Qy                    | 61     | CTGTTGAAGAGGACACCTCATCCAGTCATATTTAGAGCAAGGTCCTTGAATAGCGAG    | 120   |            |    |        |       |
| Db                    | 61     | CTGTTGAAGAGGACACCTCATCCAGTCATATTTAGAGCAAGGTCCTTGAATAGCGAG    | 120   |            |    |        |       |
| Qy                    | 121    | ATTATCGGACCGAGATACGTTGACGATACAGGATGACTACTGTGAAGATATGTTTC     | 180   |            |    |        |       |
| Db                    | 121    | ATTATCGGACCGAGATACGTTGACGATACAGGATGACTACTGTGAAGATATGTTTC     | 180   |            |    |        |       |
| Qy                    | 181    | CTAGACATTATCAGAGACATTCGAAAGCGGGTATCGAATCCACTGCAAGTAACTTCAG   | 240   |            |    |        |       |
| Db                    | 181    | CTAGACATTATCAGAGACATTCGAAAGCGGGTATCGAATCCACTGCAAGTAACTTCAG   | 240   |            |    |        |       |
| Qy                    | 241    | TCCGACGAGGAGAGCAGTCCTCTAAAAGGAGCGCAATAGACACTGTTTCAAGTCATCAGT | 300   |            |    |        |       |



| LOCUS                 | AR406007  | 2354 bp     | DNA | linear | PAT 18-DEC-2003 |
|-----------------------|---|-------------|-----|--------|-----------------|
| DEFINITION            | Sequence 1 from patent US 6630337.                                  |             |     |        |                 |
| ACCESSION             | AR406007  |             |     |        |                 |
| VERSION               | AR406007.1  | GI:40155113 |     |        |                 |
| KEYWORDS              |   |             |     |        |                 |
| SOURCE                | Unknown.  |             |     |        |                 |
| ORGANISM              | Unknown.  |             |     |        |                 |
| REFERENCE             | Unclassified.   |             |     |        |                 |
| AUTHORS               | 1 (bases 1 to 2354)   |             |     |        |                 |
| TITLE                 | Yan,C., Ye,J., Ketchum,K.A., Di Francesco,V. and Beasley,E.M.       |             |     |        |                 |
|                       | Isolated human kinase proteins, nucleic acid molecules encoding     |             |     |        |                 |
|                       | human kinase proteins, and uses thereof                             |             |     |        |                 |
| JOURNAL               | Patent: US 6630337-A 1 07-OCT-2003;                                 |             |     |        |                 |
| FEATURES              | Location/Qualifiers   |             |     |        |                 |
| source                | 1..2354   |             |     |        |                 |
|                       | /organism="unknown"   |             |     |        |                 |
| ORIGIN                | /mol_type="genomic DNA"   |             |     |        |                 |
| Query Match           | 100.0%; Score 2354; DB 6; Length 2354;                              |             |     |        |                 |
| Best Local Similarity | 100.0%; Pred. No. 0;  |             |     |        |                 |
| Matches 2354;         | Conservative 0; Mismatches 0; Indels 0; Gaps 0;                     |             |     |        |                 |
| Qy                    | 1 GCCAGCTGGGGTTACTTTTAAAAACATGCTCCATGTGCATCCCTCTTTGAAGCTTCGCACT     | 60          |     |        |                 |
| Db                    | 1 GCCAGCTGGGGTTACTTTTAAAAAACATGCTCCATGTGCATCCCTCTTTGAAGCTTCGCACT    | 60          |     |        |                 |
| Qy                    | 61 CTGTTGAAGAGGACACATCCCTCCAGTCATTATTTTAGAAGCAAGCTCTTGAATCAGGCGAG   | 120         |     |        |                 |
| Db                    | 61 CTGTTGAAGAGGACACATCCCTCCAGTCATTATTTTAGAAGCAAGCTCTTGAATCAGGCGAG   | 120         |     |        |                 |
| Qy                    | 121 ATTATCGGGACCGAGATACGTTTCACGAATAACAGGAATGACTACTGTCAAGGATATGTTTC  | 180         |     |        |                 |
| Db                    | 121 ATTATCGGGACCGAGATACGTTTCACGAATAACAGGAATGACTACTGTGAAGGATATGTTTC  | 180         |     |        |                 |
| Qy                    | 181 CTGACATTATCA CAGAGACATTGAAACGGGGTATCGAATCCACTGCAGTAAATCTTTTCAG  | 240         |     |        |                 |
| Db                    | 181 CTGACATTATCA CAGAGACATTGAAACGGGGTATCGAATCCACTGCAGTAAATCTTTTCAG  | 240         |     |        |                 |
| Qy                    | 241 TCCGCACGAGAGAACGACGCTCTTAAAGGAGCGCCAAATAGACACTGTTTCAAGTCATCAGT  | 300         |     |        |                 |
| Db                    | 241 TCCGCACGAGAGAACGACGCTCTTAAAGGAGCGCCAAATAGACACTGTTTCAAGTCATCAGT  | 300         |     |        |                 |
| Qy                    | 301 CACGTTCCGAAGAGCCACCGAAGGAAAAGATCCAGGAGTATAGAGGATCATGAGGAGGTC    | 360         |     |        |                 |
| Db                    | 301 CACGTTCCGAAGAGCCACCGAAGGAAAAGATCCAGGAGTATAGAGGATCATGAGGAGGTC    | 360         |     |        |                 |
| Qy                    | 361 ACCTGATCTGTCAAAGTGGAGACGTTTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGG  | 420         |     |        |                 |
| Db                    | 361 ACCTGATCTGTCAAAGTGGAGACGTTTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGG  | 420         |     |        |                 |
| Qy                    | 421 GTGAAGAGCGCTTTGGCAAAAGTTCTAGAGTCATTTGATCATGGCATGGATGGCATGCATG   | 480         |     |        |                 |
| Db                    | 421 GTGAAGAGCGCTTTGGCAAAAGTTCTAGAGTCATTTGATCATGGCATGGATGGCATGCATG   | 480         |     |        |                 |
| Qy                    | 481 TAGCAGTGAATAATCGTAAAAAATGTAGCCCGCTTACCCTGAAGCAGCTCGTTTCAGAAATCC | 540         |     |        |                 |
| Db                    | 481 TAGCAGTGAATAATCGTAAAAAATGTAGCCCGCTTACCCTGAAGCAGCTCGTTTCAGAAATCC | 540         |     |        |                 |
| Qy                    | 541 AAGTATTAGAGCACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGC    | 600         |     |        |                 |
| Db                    | 541 AAGTATTAGAGCACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGC    | 600         |     |        |                 |
| Qy                    | 601 TAGAATCGTTTGATCATCATGGTTCATGTTTGTATTTGTTTGAATCTACTGGCACTTAGTA   | 660         |     |        |                 |
| Db                    | 601 TAGAATCGTTTGATCATCATGGTTCATGTTTGTATTTGTTTGAATCTACTGGCACTTAGTA   | 660         |     |        |                 |
| Qy                    | 661 CTTACGATTTTCATTAAAGAAACAGCTTTCTGCCATTTCAAATTTGACCAATCAGGCAGA    | 720         |     |        |                 |
| Db                    | 661 CTTACGATTTTCATTAAAGAAACAGCTTTCTGCCATTTCAAATTTGACCAATCAGGCAGA    | 720         |     |        |                 |
| Qy                    | 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTTTACATCATATAATTAATTAACCCATACAG   | 780         |     |        |                 |



QY 1261 AGAACTGTTGACCTGGTTCGAGAAATGTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320  
Db 1261 AGAACTGTTGACCTGGTTCGAGAAATGTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320  
QY 1321 TGGATGAAGCATTGTCAGCATCTCTTCTTTCGACTTATTAAGAAAGAAATGAATGGAAATC 1380  
Db 1321 TGGATGAAGCATTGTCAGCATCTCTTCTTTCGACTTATTAAGAAAGAAATGAATGGAAATC 1380  
QY 1381 AGTGGTCTTACTATATACCTCTCTAGAAGAGATTACTTTAAGACTGTGTCACTCACTAAA 1440  
Db 1381 AGTGGTCTTACTATATACCTCTCTAGAAGAGATTACTTTAAGACTGTGTCACTCACTAAA 1440  
QY 1441 CATTCCTAATATTTTGTAAAACATTAATAATTTTGTACAGTAAGTGAATATTTGTATG 1500  
Db 1441 CATTCCTAATATTTTGTAAAACATTAATAATTTTGTACAGTAAGTGAATATTTGTATG 1500  
QY 1501 TTTTGTATCAATAGCATATAATTAACCTGTGTAGCAAGTATGCTTGTGAATGCAATGAA 1560  
Db 1501 TTTTGTATCAATAGCATATAATTAACCTGTGTAGCAAGTATGCTTGTGAATGCAATGAA 1560  
QY 1561 AAATTTAAATTAATTTTCTTTTGAATTTACCAATTTTAAATACCTTTGAAATATCCTT 1620  
Db 1561 AAATTTAAATTAATTTTCTTTTGAATTTACCAATTTTAAATACCTTTGAAATATCCTT 1620  
QY 1621 TGTGTCAGTGATAAATGTGATGATCTGTGCTTTTGTACATGAGAGTCACTCTGAAAT 1680  
Db 1621 TGTGTCAGTGATAAATGTGATGATCTGTGCTTTTGTACATGAGAGTCACTCTGAAAT 1680  
QY 1681 GATTTTCTTTCAGTAAAGGAATCTTCGACTACTTTATATCTTAAAGGAATATTCCTTA 1740  
Db 1681 GATTTTCTTTCAGTAAAGGAATCTTCGACTACTTTATATCTTAAAGGAATATTCCTTA 1740  
QY 1741 TATACCTCAAAATTTAGAACTTAACTTTTAAAGTCTTTCTTCTGTAATTTGTTGAACGGGTG 1800  
Db 1741 TATACCTCAAAATTTAGAACTTAACTTTTAAAGTCTTTCTTCTGTAATTTGTTGAACGGGTG 1800  
QY 1801 ATTATTTAATCTAGATAAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGT 1860  
Db 1801 ATTATTTAATCTAGATAAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGT 1860  
QY 1861 TAGAATCTATTAATTTTAAAGTGTGTATCTTTTCTTTCATTTGGGTGATGTCAGGTGATA 1920  
Db 1861 TAGAATCTATTAATTTTAAAGTGTGTATCTTTTCTTTCATTTGGGTGATGTCAGGTGATA 1920  
QY 1921 ACCAGACATTCATGAAAGGATGATGATGTTGTCATTTGTCAGTGTGTTTAAATAAAACC 1980  
Db 1921 ACCAGACATTCATGAAAGGATGATGATGTTGTCATTTGTCAGTGTGTTTAAATAAAACC 1980  
QY 1981 ACATACACACTTTATTTAAGATTAATAATCTAATCTGAAAGTCAAGTTTGGAAAAATGGAAT 2040  
Db 1981 ACATACACACTTTATTTAAGATTAATAATCTAATCTGAAAGTCAAGTTTGGAAAAATGGAAT 2040  
QY 2041 TTCCAGTATGTTGGTTCAGTCAGATATAAATAAGAAATCTGATGAGAGGTTTTCAG 2100  
Db 2041 TTCCAGTATGTTGGTTCAGTCAGATATAAATAAGAAATCTGATGAGAGGTTTTCAG 2100  
QY 2101 TTTTAAATACCAAGTCTTTCAGGATCTTAAATTCGCGCAGATCTGTTTATCAAAATGACA 2160  
Db 2101 TTTTAAATACCAAGTCTTTCAGGATCTTAAATTCGCGCAGATCTGTTTATCAAAATGACA 2160  
QY 2161 TAAATACGTAACCTTATAGAAATTAAGTTTATTAATTAGGCAATTTATGTCGTGATAT 2220  
Db 2161 TAAATACGTAACCTTATAGAAATTAAGTTTATTAATTAGGCAATTTATGTCGTGATAT 2220  
QY 2221 TCTTACGGGAGAAAGAGATTGTTGTAAGAGCAGTTTGGGAAGAAAGTGTCTGTAAT 2280  
Db 2221 TCTTACGGGAGAAAGAGATTGTTGTAAGAGCAGTTTGGGAAGAAAGTGTCTGTAAT 2280  
QY 2281 TTCCAGAAATTTAATTTGTTTACATAAATCTTTTGTACTTTCAGAAAAAATAAAAA 2340  
Db 2281 TTCCAGAAATTTAATTTGTTTACATAAATCTTTTGTACTTTCAGAAAAAATAAAAA 2340

QY 2341 AACAAAAAATAAAC 2354  
Db 2341 AACAAAAAATAAAC 2354  
RESULT 4  
AX392827  
LOCUS AX392827 2354 bp DNA linear PAT 23-MAR-2002  
DEFINITION Sequence 1 from Patent WO0216567.  
ACCESSION AX392827  
VERSION AX392827.1 GI:19700923  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Yan, C. C., Ye, J. C., Ketchum, K. A., di Francesco, V. C. and Beasley, E. M.  
TITLE Isolated human kinase proteins, nucleic acid molecules encoding  
human kinase proteins, and uses thereof  
JOURNAL Patent: WO 0216567-A 1 28-FEB-2002;  
Applera Corporation (US)  
FEATURES  
Location/Qualifiers  
1..2354  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 2354; DB 6; Length 2354;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCCAGCTGGGGTTACTTTAAAAAATCATGCTCATGTGATCCCTCTTGAAGCTTCGCACT 60  
Db 1 GCCAGCTGGGGTTACTTTAAAAAATCATGCTCATGTGATCCCTCTTGAAGCTTCGCACT 60  
QY 61 CTGTTGAAGAGGACACTCATCCAGTCAATTTTAAAGCAAGGTCCTTGAATGAGCGAG 120  
Db 61 CTGTTGAAGAGGACACTCATCCAGTCAATTTTAAAGCAAGGTCCTTGAATGAGCGAG 120  
QY 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTC 180  
Db 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTC 180  
QY 181 CTAGACATTTATCAGAGACATTCGAAAGCGGATTCGAAATCCACTGCACTGATTAATCTTCAG 240  
Db 181 CTAGACATTTATCAGAGACATTCGAAAGCGGATTCGAAATCCACTGCACTGATTAATCTTCAG 240  
QY 241 TCCGACGAGGAGAGCAGTCTCTAAAAGGAAAGCGCAATAGACATCTGTTCAAGTCATCAGT 300  
Db 241 TCCGACGAGGAGAGCAGTCTCTAAAAGGAAAGCGCAATAGACATCTGTTCAAGTCATCAGT 300  
QY 301 CACGTTTGAAGAGCCACCGAAGGAAAGATCCAGAGTATAGAGGATGATGAGGAGGTC 360  
Db 301 CACGTTTGAAGAGCCACCGAAGGAAAGATCCAGAGTATAGAGGATGATGAGGAGGTC 360  
QY 361 ACCTGATCTGTCAAAGTGGAGACGTTCTTAAGAGCAAGATATGAAATCGTGACACTTTGG 420  
Db 361 ACCTGATCTGTCAAAGTGGAGACGTTCTTAAGAGCAAGATATGAAATCGTGACACTTTGG 420  
QY 421 GTGAAGGAGCCTTTGGCAAAAGTTGTAGAGTCAATTTGATCATGGCATGATGGCATGATG 480  
Db 421 GTGAAGGAGCCTTTGGCAAAAGTTGTAGAGTCAATTTGATCATGGCATGATGGCATGATG 480  
QY 481 TAGCAGTGAATAATCGTAAAAAATGTAGCCGTTTACCGTGAAGCAGCTCGTTCAGAAATCC 540  
Db 481 TAGCAGTGAATAATCGTAAAAAATGTAGCCGTTTACCGTGAAGCAGCTCGTTCAGAAATCC 540  
QY 541 AAGTATTAGAGCACTTAAATAGTACTCATCCCAATAGTGTCTCCGATGTGTCCAGATGTC 600  
Db 541 AAGTATTAGAGCACTTAAATAGTACTCATCCCAATAGTGTCTCCGATGTGTCCAGATGTC 600

|    |      |        |                      |  |   |                 |
|----|------|--------|----------------------|--|---|-----------------|
| Qy | 601  | TAGAA  | TGGTTTGATCATC        | ATGGH  | CATGTTTGTGATATGTTGTGAACTACTCGGACTTAGTA    | 660             |
| Db | 601  | TAGAA  | TGGTTTGATCATCATGGT   | CATGTTTGTGATATGTTGTGAACTACTCGGACTTAGTA             | 660                                       |                 |
| Qy | 661  | CTTAG  | GAITTCATTA           | AAAGAAAACAGCTTTCTG                                 | CCATTTCAAATTTGACCA                        | CATCAGGCAGA 720 |
| Db | 661  | CTTAG  | GAITTCATTA           | AAAGAAAACAGCTTTCTG                                 | CCATTTCAAATTTGACCA                        | CATCAGGCAGA 720 |
| Qy | 721  | TGGCG  | TATCAGATCTGC         | GCAGTCAATAAATTTTTTAT                               | CATCATATAAATAAACCCATACAG                  | 780             |
| Db | 721  | TGGCG  | TATCAGATCTGC         | GCAGTCAATAAATTTTTTAT                               | CATCATATAAATAAACCCATACAG                  | 780             |
| Qy | 781  | ATCTGA | AGCCTGAAAAATATTTTGTG | TGAAGTCTGACTATGTAGTCAAAATATAAATTTCTA               | 840                                       |                 |
| Db | 781  | ATCTGA | AGCCTGAAAAATATTTTGTG | TGAAGTCTGACTATGTAGTCAAAATATAAATTTCTA               | 840                                       |                 |
| Qy | 841  | AAATGA | AAACGTGATGAAC        | GCACACTGAAAAACACAGATATCAAAGTTTGTGAC                | TTTGGAA 900                               |                 |
| Db | 841  | AAATGA | AAACGTGATGAAC        | GCACACTGAAAAACACAGATATCAAAGTTTGTGAC                | TTTGGAA 900                               |                 |
| Qy | 901  | GTGCA  | ACGTATGATGATGAAC     | ATCACA   | GTACTTTTGGTGTCTACCCGACCTACAGAGCTC 960     |                 |
| Db | 901  | GTGCA  | ACGTATGATGATGAAC     | ATCACA   | GTACTTTTGGTGTCTACCCGACCTACAGAGCTC 960     |                 |
| Qy | 961  | CCGAG  | GTCAATTTGGCTTTAG     | TTGGTCTCAGCCTTGTGATTTGGAGCATAGTTGCA 1020           |   |                 |
| Db | 961  | CCGAG  | GTCAATTTGGCTTTAG     | TTGGTCTCAGCCTTGTGATTTGGAGCATAGTTGCA 1020           |   |                 |
| Qy | 1021 | TTCTTA | TTGAATATTA           | CTTGGTTTTACAGTCTTTTACAGTCTATGATAGTAAGAGACCC 1080   |   |                 |
| Db | 1021 | TTCTTA | TTGAATATTA           | CTTGGTTTTACAGTCTTTTACAGTCTATGATAGTAAGAGACCC 1080   |   |                 |
| Qy | 1081 | TGGCA  | ATGATGGAACGAATAT     | TAGACCCATACCAACAACATGATTCAGAAAAACAAGAA 1140        |   |                 |
| Db | 1081 | TGGCA  | ATGATGGAACGAATAT     | TAGACCCATACCAACAACATGATTCAGAAAAACAAGAA 1140        |   |                 |
| Qy | 1141 | AACG   | CAAGTATTTTACCAT      | AACAGCTAGATTGGGATGAACACAGTTCTGCTGGTAGAT 1200       |   |                 |
| Db | 1141 | AACG   | CAAGTATTTTACCAT      | AACAGCTAGATTGGGATGAACACAGTTCTGCTGGTAGAT 1200       |   |                 |
| Qy | 1201 | ATGTT  | TAGAGACGTGC          | AAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAA            | CATG 1260                                 |                 |
| Db | 1201 | ATGTT  | TAGAGACGTGC          | AAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAA            | CATG 1260                                 |                 |
| Qy | 1261 | AGAA   | CTGTTTGACCTGGT       | TCGAAGATTTAGTAATATGATCCAACTCAAAGAAATTA             | CACT 1320                                 |                 |
| Db | 1261 | AGAA   | CTGTTTGACCTGGT       | TCGAAGATTTAGTAATATGATCCAACTCAAAGAAATTA             | CACT 1320                                 |                 |
| Qy | 1321 | TGGAT  | TGAAGCATTCG          | AGCATCCTTTCTTTGACTTATTTAAAAAGAAATGAAATGGGAATC 1380 |   |                 |
| Db | 1321 | TGGAT  | TGAAGCATTCG          | AGCATCCTTTCTTTGACTTATTTAAAAAGAAATGAAATGGGAATC 1380 |   |                 |
| Qy | 1381 | AGTGG  | CTTACTATATAC         | TCTCTAGAGAGATTA                                    | CTTAAAGCTGTGTCACTCAACTAAA 1440            |                 |
| Db | 1381 | AGTGG  | CTTACTATATAC         | TCTCTAGAGAGATTA                                    | CTTAAAGCTGTGTCACTCAACTAAA 1440            |                 |
| Qy | 1441 | CATT   | CTAATATTTTGT         | TAACAATTTTGTACAGTTAAGTCTAAATATTTGTATG 1500         |   |                 |
| Db | 1441 | CATT   | CTAATATTTTGT         | TAACAATTTTGTACAGTTAAGTCTAAATATTTGTATG 1500         |   |                 |
| Qy | 1501 | TTTT   | TGATCAATAG           | CATAATTAAC   | TTGTTAAGCAAGTATGCTTTGATAATGCAATTAGAA 1560 |                 |
| Db | 1501 | TTTT   | TGATCAATAG           | CATAATTAAC   | TTGTTAAGCAAGTATGCTTTGATAATGCAATTAGAA 1560 |                 |
| Qy | 1561 | AAAT   | TAAATTAATTTTCT       | TTTTTGAATTA  | CCATTTTAAATACCTTTGAAATATCCTT 1620         |                 |
| Db | 1561 | AAAT   | TAAATTAATTTTCT       | TTTTTGAATTA  | CCATTTTAAATACCTTTGAAATATCCTT 1620         |                 |
| Qy | 1621 | TGTG   | TCCAGTCA             | TAATGTGATCTTGCTTTTGTACATGAGGTC                     | CACCTCTGAAGT 1680                         |                 |
| Db | 1621 | TGTG   | TCCAGTCA             | TAATGTGATCTTGCTTTTGTACATGAGGTC                     | CACCTCTGAAGT 1680                         |                 |
| Qy | 1681 | GATTT  | TTTTTTCGATG          | AAAAAGGAATCTTCG                                    | ACTACTTATTTCTTTAAAGGAATTTCTTTA 1740       |                 |

[illegible]

```

RESULT 5
AR077239
LOCUS      2446 bp      DNA      linear
DEFINITION Sequence 9 from patent US 5962232.
ACCESSION  AR077239
VERSION     AR077239.1  GI:10003985
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 2446)
AUTHORS    Bandman,O., Hillman,J.L., Lal,P., Akerblom,I.E., Shah,P.,
            Corley,N.C. and Guegler,K.J.
TITLE       Protein kinase molecules
JOURNAL     Patent: US 5962232-A 9 05-OCT-1999;
FEATURES    Location/Qualifiers
            source          1..2446
                        /organism="unknown"
                        /mol_type="unassigned DNA"
ORIGIN

```

| ORIGIN                   |              |               |           |              |    |
|--------------------------|--------------|---------------|-----------|--------------|----|
| /mol_type=unassigned dna |              |               |           |              |    |
| Query Match              | 87.2%        | Score 2052.6; | DB 6;     | Length 2446; |    |
| Best Local Similarity    | 95.8%;       | Pred. No. 0;  |           |              |    |
| Matches 2155:            | Conservative | 0; Mismatches | 4; Indels | 90; Gaps     | 1; |



|    |      |  |      |
|----|------|--|------|
| Qy | 85   | GTCAATTAATTAGAAAGCAAGGTCCTTGGTAATGAGCGAGATTAATCGGGACCGGAGATACGTTG  | 144  |
| Db | 288  | GTCAATTAATTAGAAAGCAAGGTCCTTGGTAATGAGCGAGATTAATCGGGACCGGAGATACGTTG  | 347  |
| Qy | 145  | ACGAATACAGGAATGACTACTGTGTGAAGGATATGTTCTAGACATTAATCAGAGACATTC       | 204  |
| Db | 348  | ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTAATCAGAGACATTC         | 407  |
| Qy | 205  | AAAGCGGGTATCGAATCCACTGCAAGTAAATCTTCAGTCCGAGCAGGAGAGCAGTCCCTA       | 264  |
| Db | 408  | AAAGCGGGTATCGAATCCACTGCAAGTAAATCTTCAGTCCGAGCAGGAGAGCAGTCCCTA       | 467  |
| Qy | 265  | AAAGGAAGCGCAATAGACACTGTGTTCAAGTCATCAGTCACTGTTGCAAGAGCCACCGAAGGA    | 324  |
| Db | 468  | AAAGGAAGCGCAATAGACACTGTGTTCAAGTCATCAGTCACTGTTGCAAGAGCCACCGAAGGA    | 515  |
| Qy | 325  | AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTCGAGACG        | 384  |
| Db | 516  | -----  | 515  |
| Qy | 385  | TTCTAAGAGCAAGATATGAATCGTGGACACTTTTGGGTGAAGGAGCCCTTTGGCAAAAGTTG     | 444  |
| Db | 516  | -----AATCGTGGACACTTTTGGGTGAAGGAGCCCTTTGGCAAAAGTTG                  | 557  |
| Qy | 445  | TAGAGTGCATTGATCGGCATGAGATGGCATGCAATGTAGCAGTGAAATCGTAAAAAATG        | 504  |
| Db | 558  | TAGAGTGCATTGATCGGCATGAGATGGCATGCAATGTAGCAGTGAAATCGTAAAAAATG        | 617  |
| Qy | 505  | TAGCCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGACACTTAATAAGTA      | 564  |
| Db | 618  | TAGCCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGACACTTAATAAGTA      | 677  |
| Qy | 565  | CTGATCCCAATAGTGTCTTCGGATGTCTCCAGATGCTCAGATGCTAGAAATGTTTGAATCATCGTC | 624  |
| Db | 678  | CTGATCCCAATAGTGTCTTCGGATGTCTCCAGATGCTAGAAATGTTTGAATCATCGTC         | 737  |
| Qy | 625  | ATGTTTGTATGTGTTTGAACACTCTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA      | 684  |
| Db | 738  | ATGTTTGTATGTGTTTGAACACTCTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA      | 797  |
| Qy | 685  | GCCTTCGCCATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA       | 744  |
| Db | 798  | GCCTTCGCCATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA       | 857  |
| Qy | 745  | TAAATTTTTACATCAATAATAATTAACCCATACAGATCTGAAGCCCTGAAAAATATTTGT       | 804  |
| Db | 858  | TAAATTTTTACATCAATAATAATTAACCCATACAGATCTGAAGCCCTGAAAAATATTTGT       | 917  |
| Qy | 805  | TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCAAATGAAACGTCGTAACGCACAC        | 864  |
| Db | 918  | TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCAAATGAAACGTCGTAACGCACAC        | 977  |
| Qy | 865  | TGAAAAACACAGATATCAAAGTTGTTGCACTTTGGAAGTGCAACGTATGATGATGAACATC      | 924  |
| Db | 978  | TGAAAAACACAGATATCAAAGTTGTTGCACTTTGGAAGTGCAACGTATGATGATGAACATC      | 1037 |
| Qy | 925  | ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT       | 984  |
| Db | 1038 | ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT       | 1097 |
| Qy | 985  | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCACTTCTTATTTGAATATTAACCTTGGTT    | 1157 |
| Db | 1098 | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCACTTCTTATTTGAATATTAACCTTGGTT    | 1157 |
| Qy | 1045 | TCACAGTCTTTGAGTGTGTTTGGAGCATAGGTTGCACTTCTTATTTGAATATTAACCTTGGTT    | 1104 |
| Db | 1158 | TCACAGTCTTTGAGTGTGTTTGGAGCATAGGTTGCACTTCTTATTTGAATATTAACCTTGGTT    | 1217 |
| Qy | 1105 | GACCCATACCAACACATGATTTAGAAAAACAAGAAAAACGAAGTATTTTCAACATAACC        | 1164 |
| Db | 1218 | GACCCATACCAACACATGATTTAGAAAAACAAGAAAAACGAAGTATTTTCAACATAACC        | 1277 |
| Qy | 1165 | AGCTAGATTGGGATGAACACAGTTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT     | 1224 |



|    |      |   |      |
|----|------|---|------|
| Db | 1578 | AAATATTTTGTACAGTTAAGTGTAAATATTTGATATGTTTTGTATCAATAGCATTAATTAAC    | 1637 |
| Qy | 1525 | TTGTTAAGCAAGTATAGTCTGTGATAATGCAATAGAAAAATTTAAATTTCTTTT            | 1584 |
| Db | 1638 | TTGTTAAGCAAGTATAGTCTGTGATAATGCAATAGAAAAATTTAAATTTCTTTT            | 1697 |
| Qy | 1585 | GAAATTAACCAATTTTAAATACCTTTGAAATATCCTTTGTGCCAGTGATAAATGCGATTG      | 1644 |
| Db | 1598 | GAAATTAACCAATTTTAAATACCTTTGAAATATCCTTTGTGCCAGTGATAAATGCGATTG      | 1757 |
| Qy | 1645 | ATCTTGCCCTTTTGTACATGGAGGTCACTCTGGAAGTGATTTTTTTTTCAGTAAAGGAAAT     | 1704 |
| Db | 1758 | ATCTTGCCCTTTTGTACATGGAGGTCACTCTGGAAGTGATTTTTTTTTCAGTAAAGGAAAT     | 1817 |
| Qy | 1705 | CTTGACACTACTTATATTTCTTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAAC | 1764 |
| Db | 1818 | CTTGACACTACTTATATTTCTTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAAC | 1877 |
| Qy | 1765 | TTTAAAAAGTTTTTCTCTGTAAATTTGTTGAAACGGGTGATTAATTAATTAACCTAGATAAGCA  | 1824 |
| Db | 1878 | TTTAAAAAGTTTTTCTCTGTAAATTTGTTGAAACGGGTGATTAATTAATTAACCTAGATAAGCA  | 1937 |
| Qy | 1825 | GGTACTAGAAACCAAAAACCTCAGAAAAATGTTTACTGTGTAGAAATCTATTAATTTTAAAGTG  | 1884 |
| Db | 1938 | GGTACTAGAAACCAAAAACCTCAGAAAAATGTTTACTGTGTAGAAATCTATTAATTTTAAAGTG  | 1997 |
| Qy | 1885 | TTGTATTTCTTTTTCATTTGGGTGATGTCAGGGTGATAACACAGACATTCATCGGAAGGCATG   | 1944 |
| Db | 1998 | TTGTATTTCTTTTTCATTTGGGTGATGTCAGGGTGATAACACAGACATTCATCGGAAGGCATG   | 2057 |
| Qy | 1945 | CAGTTTTGTCCTATGTGACAGTTGTGTTTAAATAAAACCAATACACACATTTATTTAAAGATTA  | 2004 |
| Db | 2058 | CAGTTTTGTCCTATGTGACAGTTGTGTTTAAATAAAACCAATACACACATTTATTTAAAGATTA  | 2117 |
| Qy | 2005 | AAATCTAACTGGAAAAGTCAGCTTGGA AAAATGGACATTTCCAAGTATGTTTGGTGAAGTCAC  | 2064 |
| Db | 2118 | AAATCTAACTGGAAAAGTCAGCTTGGA AAAATGGACATTTCCAAGTATGTTTGGTGAAGTCAC  | 2177 |
| Qy | 2065 | AGATATAAAAATAGAAAATCTCGATGACAGGTTTTTCAGTTTTTAAATACCAAGTCCTTAGGAG  | 2124 |
| Db | 2178 | AGATATAAAAATAGAAAATCTCGATGACAGGTTTTTCAGTTTTTAAATACCAAGTCCTTAGGAG  | 2237 |
| Qy | 2125 | TCCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTATAAGAAAT   | 2184 |
| Db | 2238 | TCCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTATAAGAAAT   | 2297 |
| Qy | 2185 | AAGTTTATTAATTAGGCAATTTATGTCGTGTGATAATCTTTACGGGAGAAAGAGGATTTGA     | 2244 |
| Db | 2298 | AAGTTTATTAATTAGGCAATTTATGTCGTGTGATAATCTTTACGGGAGAAAGAGGATTTGA     | 2357 |
| Qy | 2245 | TTGGAAAGCAGTTTGGGAAGAAAGTGCCTGCTGAAATTTCCAGAAATTTAAATTTGATTCGTTA  | 2304 |
| Db | 2358 | TTGGAAAGCAGTTTGGGAAGAAAGTGCCTGCTGAAATTTCCAGAAATTTAAATTTGATTCGTTA  | 2417 |
| Qy | 2305 | CATAAACTTTTTTGACTTTCAGAAAAAAA                                     | 2333 |
| Db | 2418 | CATAAACTTTTTTGACTTTCAGAAAAAAA                                     | 2446 |

|            |   |   |     |        |                 |
|------------|---|---|-----|--------|-----------------|
| RESULT 7   |   |   |     |        |                 |
| AX056410   |   |   |     |        |                 |
| LOCUS      | AX056410  | 2488 bp   | DNA | linear | PAT 13-JAN-2001 |
| DEFINITION | Sequence 54 from Patent WO0073469.                                |   |     |        |                 |
| ACCESSION  | AX056410  |   |     |        |                 |
| VERSION    | AX056410.1  | GI:12229117   |     |        |                 |
| KEYWORDS   | .   |   |     |        |                 |
| SOURCE     | Homo sapiens (human)  |   |     |        |                 |
| ORGANISM   | Homo sapiens  |   |     |        |                 |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |   |     |        |                 |
|            | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.        |   |     |        |                 |
| REFERENCE  | 1   | Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S. |     |        |                 |
| AUTHORS    | TITLE   |   |     |        |                 |
| TITLE      | Protein kinases   |   |     |        |                 |



QY 85 GTCAATATTAGAACCAAGTCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 144  
Db 161 GTCAATATTAGAACCAAGTCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 220  
QY 145 ACGAATACAGGAATGACATCTAGTGTGAAGGATATGTTCTTAGACATTAATCACAGAGACATTG 204  
Db 221 ACGAATACAGGAATGACATCTAGTGTGAAGGATATGTTCTTAGACATTAATCACAGAGACATTG 280  
QY 205 AAAGCGGTATCGAATCCACTGCACTAAATCTTCAAGTCCGAGCAGGAGAAAGCAGTCTCTA 264  
Db 281 AAAGCGGTATCGAATCCACTGCACTAAATCTTCAAGTCCGAGCAGGAGAAAGCAGTCTCTA 340  
QY 265 AAAGGAAGCGCAATAGACACATCTGTTCAAGTCATCAGTCAGTTCGAGAGCCACCGAAGCA 324  
Db 341 AAAGGAAGCGCAATAGACACATCTGTTCAAGTCATCAGTCAGTTCGAGAGCCACCGAAGCA 400  
QY 325 AAAGATCCAGAGATATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAAGTGAGAGCG 384  
Db 401 AAAGATCCAGAGATATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAAGTGAGAGCG 460  
QY 385 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGCCCTTTGGCAAAAGTTG 444  
Db 461 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGCCCTTTGGCAAAAGTTG 520  
QY 445 TAGAGTCATTCATCGCATGCGATGCGATGTCATGTCAGTGTGAGCAATCGTAAAGATG 504  
Db 521 TAGAGTCATTCATCGCATGCGATGCGATGTCATGTCAGTGTGAGCAATCGTAAAGATG 580  
QY 505 TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATAGAGCAGCTTAAATAGTA 564  
Db 581 TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATAGAGCAGCTTAAATAGTA 640  
QY 565 CTGATCCCAATAGTGTCTCCGATGTCGAGATGTCAGATGCTAGAATGTTGATCATCATGTC 624  
Db 641 CTGATCCCAATAGTGTCTCTCCGATGTCGAGATGTCAGATGCTAGAATGTTGATCATCATGTC 700  
QY 625 ATGTTGTATGTTGTTGAACTAGTGGGACTTAGTACTTACATTCATTAAGAAACACA 684  
Db 701 ATGTTGTATGTTGTTGAACTAGTGGGACTTAGTACTTACATTCATTAAGAAACACA 760  
QY 685 GCTTTCTGCCATTTCAAAATGACACATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 744  
Db 761 GCTTTCTGCCATTTCAAAATGACACATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 820  
QY 745 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 804  
Db 821 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 880  
QY 805 TTGTGAAGTCTGACTATGATGATCAATATTAATTTCAAAATGAAGCTGATGAACGACAC 864  
Db 881 TTGTGAAGTCTGACTATGATGATCAATATTAATTTCAAAATGAAGCTGATGAACGACAC 940  
QY 865 TGAAAAACACAGATATCAAGTTGTTGACTTTGGAGTGCAACGATGATGATGAACATC 924  
Db 941 TGAAAAACACAGATATCAAGTTGTTGACTTTGGAGTGCAACGATGATGATGAACATC 1000  
QY 925 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGGTT 984  
Db 1001 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGGTT 1060  
QY 985 GGTCTCAGCTTTGATGTTGGAGCATAGTTGATTTCTTATTTGAATATTAACCTTGGTT 1044  
Db 1061 GGTCTCAGCTTTGATGTTGGAGCATAGTTGATTTCTTATTTGAATATTAACCTTGGTT 1120  
QY 1045 TCACAGTCTTCAGACTCATGATGATGAAGACCTCGCATGATGCAAGCAATATTAG 1104  
Db 1121 TCACAGTCTTCAGACTCATGATGATGAAGACCTCGCATGATGCAAGCAATATTAG 1180  
QY 1105 GACCCATACCAACACATGATTTAGAAAAACGAAAGGATTTTACCATAACC 1164  
Db 1181 GACCCATACCAACACATGATTTAGAAAAACGAAAGGATTTTACCATAACC 1240  
QY 1165 AGCTAGATTGGGATGAACACAGTTCTGCTGTGTAGATATGTTAGGAGACGCTGCAAAACCGT 1224

Db 1241 AGCTAGATTGGGATGAACACAGTTCTGCTGTGTAGATATGTTAGGAGACGCTGCAAAACCGT 1300  
QY 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAACTGAGAAACTGTTTGACCTGGTTCGAA 1284  
Db 1301 TGAAGGAATTTATGCTTTGTCATGATGAAGAACTGAGAAACTGTTTGACCTGGTTCGAA 1360  
QY 1285 GAATGTTAGAAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCGAGCATCCTT 1344  
Db 1361 GAATGTTAGAAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCGAGCATCCTT 1420  
QY 1345 TCTTTGACTTATTAAGAAAGAAATGAATGGAATCAGTGGTCTTACTATATACCTTCT 1404  
Db 1421 TCTTTGACTTATTAAGAAAGAAATGAATGGAATCAGTGGTCTTACTATATACCTTCT 1480  
QY 1405 AGAAGAGATTACTTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1464  
Db 1481 AGAAGAGATTACTTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1540  
QY 1465 AAATATTTTGTGATCAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC 1524  
Db 1541 AAATATTTTGTGATCAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC 1600  
QY 1525 TTGTTAAGCAAGTATGTTCTTGAATGCAATAGAAATTAATAATTTTCTTTT 1584  
Db 1601 TTGTTAAGCAAGTATGTTCTTGAATGCAATAGAAATTAATAATTTTCTTTT 1660  
QY 1585 GAAATTTACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGTGATTG 1644  
Db 1661 GAAATTTACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGTGATTG 1720  
QY 1645 ATCTTGCTTTTGTACATGAGGAGTCACTCTGAAAGTGAATTTTGTGATGAAGGAAT 1704  
Db 1721 ATCTTGCTTTTGTACATGAGGAGTCACTCTGAAAGTGAATTTTGTGATGAAGGAAT 1780  
QY 1705 CTTGACTACTTTTATATTTCTTAAAGGAATA 1733  
Db 1781 CTTGACTACTTTTAAATTTTAAATTTTAAATTTTAAATTTTCTTTT 1809

RESULT 9  
AX961896 1792 bp DNA linear PAT 14-JAN-2004  
LOCUS Sequence 107 from Patent WO03104277.  
DEFINITION AX961896  
ACCESSION AX961896  
VERSION AX961896.1 GI:40881306  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.  
Stat6 activation gene  
Patent: WO 03104277-A 107 18-DEC-2003;  
Asahi Kasei Kabushiki Kaisha (JP)  
Location/Qualifiers  
FEATURES  
source  
1. 1792  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
CDS  
1. 1446  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAF06457.1"  
/db\_xref="GI:40881307"  
/translation="MRHSKRTHYCDPWSRESNGHESYRSGSHKRRKRSHSSTQENRHCK  
PHQFESDCHYLEARSLSNRDHYDRYNDYCEGYPRHYHRDIESGYRHHCS  
KSSVRSRRSPKRRNRHCSHQSRKRSRSEDDEEGLHCOSGVLRLARE  
IVDTLGBGAFKVVCEIDHGMGMHVAKIVKNGRYREARSIVQLVHLNSDPNS  
VFERCVQLEWFDHGHGVCIVPELGLSTYDFIKENSLPFOIDHROMAYOICOSINF  
LHNKLTDLKPENILFKVSDYVKKNSKMKRDERILKNWTDIKVDFSGSATYDDEHH  
STLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYLGFTVFQTHDSKHLAMMERI

|                            |   |  |
|----------------------------|---|--|
| ORIGIN                     | LGPIPOHMIQKTRKRYFHNNQDWDHSSAGRYVRRRCKPLKEFMLCHDEHEKLPD<br>LVRRMLEYDPTQRTILDEALQHPFFDLKKK" |  |
| Query Match                | 69.2%; Score 1629.4; DB 6; Length 1792;   |  |
| Best Local Similarity      | 99.9%; Pred. No. 9.1e-290;  |  |
| Matches 1630; Conservative | 0; Mismatches 2; Indels 0; Gaps 0;  |  |
| QY                         | 85 GTCATTATTTAGAACGAGTCCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG 144                        |  |
| Db                         | 161 GTCATTATTTAGAACGAGTCCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG 220                       |  |
| QY                         | 145 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCCCTAGACATATATCACAGACATTTG 204                     |  |
| Db                         | 221 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCCCTAGACATATATCACAGACATTTG 280                     |  |
| QY                         | 205 AAAGGGGTATCGAATCCACTGCGAGTAATCTTCAGTCCGCGAGGAGAGAGTCCCTA 264                          |  |
| Db                         | 281 AAAGGGGTATCGAATCCACTGCGAGTAATCTTCAGTCCGCGAGGAGAGAGTCCCTA 340                          |  |
| QY                         | 265 AAAGGAGCGCAATAGACACTGTTCAAGTCATCACTGTCGAAAGAGCCACCGAAGGA 324                          |  |
| Db                         | 341 AAAGGAGCGCAATAGACACTGTTCAAGTCATCACTGTCGAAAGAGCCACCGAAGGA 400                          |  |
| QY                         | 325 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAAAAGTCGAGACG 384                       |  |
| Db                         | 401 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAAAAGTCGAGACG 460                       |  |
| QY                         | 395 TTCTAAGAGCAAGATATGAATCTGGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 444                        |  |
| Db                         | 461 TTCTAAGAGCAAGATATGAATCTGGGTGAAGGAGCCCTTTGGCAAGTTG 520                                 |  |
| QY                         | 445 TAGAGTGCAATGATCGGATGATGCGATGTCATGATGATGATGATGATGATGATGATGATG 504                      |  |
| Db                         | 521 TAGAGTGCAATGATCGGATGATGCGATGTCATGATGATGATGATGATGATGATGATGATG 580                      |  |
| QY                         | 505 TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATAGAGCACTTAAATAGTA 564                       |  |
| Db                         | 581 TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATAGAGCACTTAAATAGTA 640                       |  |
| QY                         | 565 CTGATCCCAATAGTCTCTCGATGTCGAGATGCTAGATGCTAGATGCTAGATGCTAGATGCT 624                     |  |
| Db                         | 641 CTGATCCCAATAGTCTCTCGATGTCGAGATGCTAGATGCTAGATGCTAGATGCTAGATGCT 700                     |  |
| QY                         | 625 ATGTTGTATGTTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 684                         |  |
| Db                         | 701 ATGTTGTATGTTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 760                         |  |
| QY                         | 685 GCTTCTGCCATTTCAAAATGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 744                         |  |
| Db                         | 761 GCTTCTGCCATTTCAAAATGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 820                         |  |
| QY                         | 745 TAAATTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCCTGAAATATTTGT 804                        |  |
| Db                         | 821 TAAATTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCCTGAAATATTTGT 880                        |  |
| QY                         | 805 TTGTGAAGTCTGACTATGATGATCAATATAATTTCTTAAATGAAACGTCATGAAGCAGC 864                       |  |
| Db                         | 881 TTGTGAAGTCTGACTATGATGATCAATATAATTTCTTAAATGAAACGTCATGAAGCAGC 940                       |  |
| QY                         | 865 TGAATAAACACAGATATCAAAAGTTGTTGACTTTGGAGTGCAACGTCATGATGATGAACATC 924                    |  |
| Db                         | 941 TGAATAAACACAGATATCAAAAGTTGTTGACTTTGGAGTGCAACGTCATGATGATGAACATC 1000                   |  |
| QY                         | 925 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGTCACTTTTGGCTTTAGGTT 984                        |  |
| Db                         | 1001 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGTCACTTTTGGCTTTAGGTT 1060                      |  |
| QY                         | 985 GGTCTCAGCCTTTGATGTTTGGAGCATAGTTGTCATTTCTTATTTGAATATTAACCTTTGTT 1044                   |  |
| Db                         | 1061 GGTCTCAGCCTTTGATGTTTGGAGCATAGTTGTCATTTCTTATTTGAATATTAACCTTTGTT 1120                  |  |
| QY                         | 1045 TCACAGTCTTTTCAGACTCATGATGATGAAGAGCAGCTCGCAATGATGCAATATATTAG 1104                     |  |

|            |  |  |
|------------|--|--|
| Db         | 1121 TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGGAAACGAATATTAG 1180   |  |
| QY         | 1105 GACCCATACCACAAACATGATTTCAGAAAAACAAGAAAAACGCAAGTATTTTTCACCAATAACC 1164   |  |
| Db         | 1181 GACCCATACCACAAACATGATTTCAGAAAAACAAGAAAAACGCAAGTATTTTTCACCAATAACC 1240   |  |
| QY         | 1165 AGCTAGATTGGGATGAACACAGTTCCTGCTGGTATATATGTTAGGAGACGCTGCAAAACCGT 1224   |  |
| Db         | 1241 AGCTAGATTGGGATGAACACAGTTCCTGCTGGTATATATGTTAGGAGACGCTGCAAAACCGT 1300   |  |
| QY         | 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTGGACCTGGTTCGAA 1284   |  |
| Db         | 1301 TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTGGACCTGGTTCGAA 1360   |  |
| QY         | 1285 GAATGTTTAGAATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCAATTCGAGCATCCTT 1344   |  |
| Db         | 1361 GAATGTTTAGAATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCAATTCGAGCATCCTT 1420   |  |
| QY         | 1345 TCTTTGACTTTTAAAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTCT 1404  |  |
| Db         | 1421 TCTTTGACTTTTAAAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTCT 1480  |  |
| QY         | 1405 AGAAGAGATTACTTAAAGACTGTGTCAGTCAACTTAAACATTTCTAATATTTTGTAAACATT 1464   |  |
| Db         | 1481 AGAAGAGATTACTTAAAGACTGTGTCAGTCAACTTAAACATTTCTAATATTTTGTAAACATT 1540   |  |
| QY         | 1465 AAATATTTTGTACAGTTAAGTGTAAATATTTGATGTTTGTATCAATCAATGATCAATTAAC 1524  |  |
| Db         | 1541 AAATATTTTGTACAGTTAAGTGTAAATATTTGATGTTTGTATCAATCAATGATCAATTAAC 1600  |  |
| QY         | 1525 TTGTTAAGCAAGTATGCTTGTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1584   |  |
| Db         | 1601 TTGTTAAGCAAGTATGCTTGTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1660   |  |
| QY         | 1585 GAAATACCAATTTTAAATACCTTTGAAATATCTTGTGTCAGTCAATGCAATGCAATGCAAT 1644  |  |
| Db         | 1661 GAAATACCAATTTTAAATACCTTTGAAATATCTTGTGTCAGTCAATGCAATGCAATGCAAT 1720  |  |
| QY         | 1645 ATCTTGCTTTTGTACATGAGGTCACCTCTGAAGTCAATTTTGTGAGTAAAGGAAAT 1704   |  |
| Db         | 1721 ATCTTGCTTTTGTACATGAGGTCACCTCTGAAGTCAATTTTGTGAGTAAAGGAAAT 1780   |  |
| QY         | 1705 CTGACTACTTTT 1716   |  |
| Db         | 1781 CTGACTACTTTT 1792   |  |
| RESULT 10  |  |  |
| BC063116   |  |  |
| LOCUS      | Homo sapiens CDC-like kinase 4, mRNA (cdna clone IMAGE:30324121),  |  |
| DEFINITION | partial cds.   |  |
| ACCESSION  | BC063116   |  |
| VERSION    | BC063116.1 GI:38970005   |  |
| KEYWORDS   |  |  |
| SOURCE     | Homo sapiens (human)   |  |
| ORGANISM   |  |  |
| REFERENCE  |  |  |
| AUTHORS    |  |  |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 3492)   |  |
|            | Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Hsieh, F., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.F., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullighy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.C., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., |  |

Pahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A.C., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1' to 3492)

Strausberg, R.

Direct Submission

Submitted (02-DEC-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: GPCs help desk

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzyzinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nee, Pawan Pandoh, Anna-Lisa Prabhu, Farvaneh Saeedi, Jacqueline Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 51 Row: a Column: 9

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10190705.

FEATURES

source

1..3492

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30324121"

/tissue\_type="Skin and meninges pool- skin, dura matter, pia matter, choroid plexus."

/clone\_lib="NIH MGC\_186"

/lab\_host="DH10B"

/note="Vector: pDNR-LIB"

ORIGIN

Query Match 56.0%; Score 1318.4; DB 9; Length 3492;

Best Local Similarity 99.2%; Pred. No. 1.3e-232;

Matches 1325; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 398 ATATCAATCGTGGACACTTTGGGTGAAGGAGCCCTTCGCAAGTGTAGAGTCGATTGA 457

DB 2149 AGATGAATCGTGGACACTTTGGGTGAAGGAGCCCTTCGCAAGTGTAGAGTCGATTGA 2208

QY 458 TCATGGCATGGATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATGTAGGCCGTTACCG 517

DB 2209 TCATGGCATGGATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATGTAGGCCGTTACCG 2268

QY 518 TGAAGCAGCTCGTTCAGAAATCCAAAGTATAGAGCACTTAATAGTACTGTATCCCAATAG 577

DB 2269 TGAAGCAGCTCGTTCAGAAATCCAAAGTATAGAGCACTTAATAGTACTGTATCCCAATAG 2328

QY 578 TGTCTTCGATGTGCCAGATGCTAGATGGTTTTCATCATGTGTCATGTTTGTATGTT 637

DB 2329 TGTCTTCGATGTGCCAGATGCTAGATGGTTTTCATCATGTGTCATGTTTGTATGTT 2388

QY 638 GTTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACACAGCTTCTGCCATT 697

DB 2389 GTTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACACAGCTTCTGCCATT 2448

QY 698 TCAAAATTCAGCACATCAGGACAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA 757

DB 2449 TCAAAATTCAGCACATCAGGACAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA 2508

QY 758 TCATAATAAATTAACCCCATACAGATCTGAAGCCTGAAATATTTTTTGTGGAAGTCTGA 817

DB 2509 TCATAATAAATTAACCCCATACAGATCTGAAGCCTGAAATATTTTTTGTGGAAGTCTGA 2568

QY 818 CTATGTAGTCAAAATATAATTTCTAAATGAACGATGAACGCCACACTGAAACACACAGA 877

DB 2569 CTATGTAGTCAAAATATAATTTCTAAATGAACGATGAACGCCACACTGAAACACACAGA 2628

QY 878 TATCAAAAGTCTGTGACTTTTGAAGTGAACGCTATGATGATGAACATCAAGTACTTTGGT 937

DB 2629 TATCAAAAGTCTGTGACTTTTGAAGTGAACGCTATGATGATGAACATCAAGTACTTTGGT 2688

QY 938 GTCTACCCGGCAGTACAGAGCTCCGAGGTCAATTTTGGCTTTAGGTTGGTCTCAGCCCTTG 997

DB 2689 GTCTACCCGGCAGTACAGAGCTCCGAGGTCAATTTTGGCTTTAGGTTGGTCTCAGCCCTTG 2748

QY 998 TGATGTTTGGAGCATAGGTTGCATTTCTTATTGAATATTACCTTGGTTTTCAGAGTCTTTCA 1057

DB 2749 TGATGTTTGGAGCATAGGTTGCATTTCTTATTGAATATTACCTTGGTTTTCAGAGTCTTTCA 2808

QY 1058 GACTCATGATAGTAAAGAGCAGTCCGCAATGATGGAAGCAATATTAGGACCCATACACACA 1117

DB 2809 GACTCATGATAGTAAAGAGCAGTCCGCAATGATGGAAGCAATATTAGGACCCATACACACA 2868

QY 1118 ACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCACCATTAACAGTACTAGTTGGA 1177

DB 2869 ACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCACCATTAACAGTACTAGTTGGA 2928

QY 1178 TGAACACAGTCTGCTGGTGTAGATGTTTGAAGAGCGCTGCAACCGTTGAAGGAATTTAT 1237

DB 2929 TGAACACAGTCTGCTGGTGTAGATGTTTGAAGAGCGCTGCAACCGTTGAAGGAATTTAT 2988

QY 1238 GCTTTGTCATGATGAAGAAACATGAGAACTGTTTGACCTGGTTGCAAGCAATGTTAGAATA 1297

DB 2989 GCTTTGTCATGATGAAGAAACATGAGAACTGTTTGACCTGGTTGCAAGCAATGTTAGAATA 3048

QY 1298 TGATCCAACTCAAAGAATTACTTTGGATGAAGCATTGAGCATCTCTTTTGTGACTTATT 1357

DB 3049 TGATCCAACTCAAAGAATTACTTTGGATGAAGCATTGAGCATCTCTTTTGTGACTTATT 3108

QY 1358 AAAAAAGAAATGAATGGGAATCAGTGGTCTTATATATATCTTCTAGAGAGAGATTACT 1417

DB 3109 AAAAAAGAAATGAATGGGAATCAGTGGTCTTATATATCTTCTAGAGAGAGATTACT 3168

QY 1418 TAAGACTGTGTCAGTCAACTAAACATTTCTATATTTTTTGTAAACATTAATTTTGTGA 1477

DB 3169 TAAGACTGTGTCAGTCAACTAAACATTTCTATATTTTTTGTAAACATTAATTTTGTGA 3228

QY 1478 CAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATAAATAACTTGTGTAAGCAAGT 1537

DB 3229 CAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATAAATAACTTGTGTAAGCAAGT 3288

QY 1538 ATGCTCTTGAATGCAATAGAAAAATTAATAATTTTTTCTTTTGAATTTACCAATTT 1597

DB 3289 ATGCTCTTGAATGCAATAGAAAAATTAATAATTTTTTCTTTTGAATTTACCAATTT 3348

QY 1598 TTAATACCTTTGAAATATCTTTGTGTCAGTGAATAATGTGATGATCTTGCCTTTTG 1657

DB 3349 TTAATACCTTTGAAATATCTTTGTGTCAGTGAATAATGTGATGATCTTGCCTTTTG 3408

QY 1658 TACATGGAGGTCACTCTGAAAGTGAATTTTTTTTGTAGTAAAGGAATCTTCTACTACTTTA 1717

DB 3409 TACATGGAGGTCACTCTGAAAGTGAATTTTTTTTGTAGTAAAGGAATCTTCTACTACTTTA 3468

QY 1718 TATTCTTAAAGGAATA 1733









|    |      |   |      |
|----|------|---|------|
| Db | 681  | TAGGCCGTTACCGGGAGGCGAGCTCGTTCTCGAAATCCAAGTATTGGAGCATTTGGAACAGCA   | 740  |
| Qy | 565  | CTGATCCCAATPAGTGTCTTCCGATGTGTCACAGATGCTAGAAATGGTTTGGATCATCATGCGTC | 624  |
| Db | 741  | CTGACCCCAACAGTGTCTTCCGATGCGTCACAGATGCTAGAGTGGTTTGGATCATCATGCGTC   | 800  |
| Qy | 625  | ATGTTTGTATTTGTGTTTGAACCTACTGGGACTTTAGTACTTACGATTTCAATTAAAGAAAACA  | 684  |
| Db | 801  | ATGTTTGTATTTGTGTTTGGAGCTGTGGGACTTTAGTACTATGATTTTATTTAAAGAAAATA    | 860  |
| Qy | 685  | GCTTTCTGCCATTTCAATTTGACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA        | 744  |
| Db | 861  | GTTTTCTGCCATTTCAATTTGATCACATCAGGCAAAATGGCTTATCAGATCTGCCAGTCTA     | 920  |
| Qy | 745  | TAAATTTTTTACATCATAAATAAAATTAACCOATACAGATCTGAAGCCTGAAAAATATTTTGT   | 804  |
| Db | 921  | TAAATTTTTTACATCATATAATAATTAAACACACGGACCTAANAACCTGAANAATATTTTAT    | 980  |
| Qy | 805  | TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAATGAAACGTGATGAACGCACAC   | 864  |
| Db | 981  | TTGTGAAGTCTGACTATGTAGTCAAAATACAATTTCTAAAAATGAAACGAGATGAGCGCACAT   | 1040 |
| Qy | 865  | TGAAAAACACAGATATCAAGTGTGTTCGACTTTGGAAGTGCACGATATGATGAACATC        | 924  |
| Db | 1041 | TGAAAAACACAGATATCAAAAGTTTGTGATTTTGGAAATGCAACATATGACGACGAACATC     | 1100 |
| Qy | 925  | ACAGTACTTTGGTGTCTTACCCGGCACATACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT    | 984  |
| Db | 1101 | ATAGTACTTTTGGTGTCCAAAGGCACATACAGGGCTCCAGAGGTCATTTTGGCTCTAGGTT     | 1160 |
| Qy | 985  | GGTCTCAGCCCTTGTGATGTTTGGAGCATPAGTGTGCATTTCTTATTGAAATATTACCTTGGTT  | 1044 |
| Db | 1161 | GGTCTCAGCCCTTGTGATGTTTGGAGCATAGSCTGCAITTTCTATTGAGTACTACCTTGGGT    | 1220 |
| Qy | 1045 | TCACAGTCTTTTCAGACTCATGATAGTAAAGAGACCTTGGCAATGATGGAACGAATATTAG     | 1104 |
| Db | 1221 | TCACAGTCTTTTCAGACCCACGATAGTAAAGAGCACCTTGGCAATGATGGAACGCGATCTTAG   | 1280 |
| Qy | 1105 | GACCCATACCAACACATGATTTTCAGAAAAACAAGAAAAACGCAAGTATTTTTCACCATAACC   | 1164 |
| Db | 1281 | GACCCATCCAGCAATATGATTCAGAAAGACAGGAAACGCAAGTATTTTCCACATNACC        | 1340 |
| Qy | 1165 | AGCTAGATTGGGATGAACACAGTTTCTGCTGTGTAGATATGTTTGGAGACGCTGCAAAACCGT   | 1224 |
| Db | 1341 | AGCTAGATTGGGACGAGCATAGTTTCAGCTGGAGATATGTTTGGAGACGCTGCAAAACCGT     | 1400 |
| Qy | 1225 | TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGATGAAGAACTGTTTGAACCTGGTTGCA   | 1284 |
| Db | 1401 | TAAAGGAATTTATGCTGTGTCTATGACGAAGAGCATGAGAAGCTGTTTGGACCTGGTTTCCAA   | 1460 |
| Qy | 1285 | GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCGACGATCCTT     | 1344 |
| Db | 1461 | GAATGTTGGAGTATGACCCGAGAGAGGATTCACCTTGGATGAAGCAATTCGACGACCCCTT     | 1520 |
| Qy | 1345 | TCCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTTACTATATCTTCTCT     | 1404 |
| Db | 1521 | TCCTTTGACTTATTAAGAAAGAAATG- AGTGGAGTCAAGGGTCTTCTGTGTACTTCTCTA     | 1579 |
| Qy | 1405 | AGAAGAGATTACTTAAAGACTGTGCAGTCAACTAAAACTTAATAATTTTGTGTAACAT        | 1464 |
| Db | 1580 | GGAGCAGTTACTTCCAGACTGTGTCACTCAACTAAACCTTCTAATAATTTTGTGTAACAT      | 1639 |
| Qy | 1465 | AAATATTTTGTACAGTTTAAAGTGAATATGATGTTTGTATCAATACGATAATTAAC          | 1524 |
| Db | 1640 | AAATATTTTGTACAGTTTAAATGAAATGAAATGATGTTTGTATCAATAGCAATTAAC         | 1699 |
| Qy | 1525 | TTGTTTAAAGCAAGTATGGTCTTGTAAATGCAATTAGAAAAATTAATAATTTTCTTTT        | 1583 |
| Db | 1700 | TTGTTTAAAGTATGGTGTGATTAATGAAATGCAATGCAATGCAATGCAATTTTCTTTT        | 1759 |
| Qy | 1584 | -TGAATTPACATTTTAAATACTTTTGAATAATCTTTTGTGTGTCAGTGAATAATGTGAT       | 1642 |
| Db | 1760 | ATGTTTAAATGGCGCTTTAAAGGCTTTTAAAGATGCGCTTTGTCAGTGAATAATGTGAT       | 1819 |

|                       |   |   |                            |
|-----------------------|---|---|----------------------------|
| Qy                    | 1643  | TGATCTTGGCTTTTGTACATGAGAGTTCACCTCTGAAGTGAATTTTTTTTGTAGTAAAAGGAA | 1702                       |
| Db                    | 1820  | TGCTCCCATCTTTTGTACATGAAGGTTGACTCTGAAGTGAATTTTTTTTCTCAAGTAA      | 1879                       |
| Qy                    | 1703  | A   | 1703                       |
| Db                    | 1880  | A   | 1880                       |
| RESULT 13             |   |   |                            |
| AX961894              |   |   |                            |
| LOCUS                 | AX961894  | 1881 bp   | DNA linear PAT 14-JAN-2004 |
| DEFINITION            | Sequence 105 from Patent WO03104277.  |   |                            |
| ACCESSION             | AX961894  |   |                            |
| VERSION               | AX961894.1  | GI:40881304   |                            |
| KEYWORDS              | Mus musculus (house mouse)  |   |                            |
| SOURCE                | Mus musculus  |   |                            |
| ORGANISM              | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;  |   |                            |
| REFERENCE             | 1   |   |                            |
| AUTHORS               | Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.  |   |                            |
| TITLE                 | Stat6 activation gene   |   |                            |
| JOURNAL               | Patent: WO 03104277-A 105 18-DEC-2003;  |   |                            |
|                       | Asahi Kasei Kabushiki Kaisha (JP)   |   |                            |
| FEATURES              | Location/Qualifiers   |   |                            |
| source                | 1. 1881   |   |                            |
|                       | /organism="Mus musculus"  |   |                            |
|                       | /mol_type="unassigned DNA"  |   |                            |
|                       | /db_xref="taxon:10090"  |   |                            |
|                       | 90..1535  |   |                            |
|                       | /note="unnamed protein product"   |   |                            |
|                       | /codon_start=1  |   |                            |
|                       | /protein_id="CAF06456.1"  |   |                            |
|                       | /db_xref="GI:40881305"  |   |                            |
|                       | /translation="MRHSKRTCPDWDSRESWGSHESYSGHKKRSHSHSTQENRRHCK<br>PHHOFKSDCHVLEAECLNERDYDRRYIDEVENDYCEGVPRPHRDVESTYIRVICS<br>KSVRRSRSPKRRNRPNCASHOSKSHRRKRSRSEDDEGHLICQSGDLVRAVE<br>IVDTLGEAGPKVVECDIGMDGGLHVAKIVKNVGRZARSEIQVLEHNSHTDENS<br>VPRCQVLEMDHGHVCIYFELLGLSYDFIKENSFLPKNDIKVDFQSGATYDDHH<br>LHNKLTHTDLKPELTVKSYDVVYVYNSKMKRDERTLKNVDIKVDFQSGATYDDHH<br>SLTVPVRYRPAVPIVLAIGWSQPCDVMSIGCILIEYLGTFVQFTHDSKEHLAMMRI<br>LGPVPAHMIQKTRKYRHYHNOLDWDEHSSAGRYVRRCKPLKFLCHDEEHEKLPD<br>LVRRMLEYDPAIRITLDEALQHPFDLLKRX" |   |                            |
| CDS                   |   |   |                            |
| ORIGIN                |   |   |                            |
| Query Match           | 54.3%;  | Score 1277.4;   | DB 6; Length 1881;         |
| Best Local Similarity | 87.9%;  | Pred. No. 5.1e-225;   |                            |
| Matches 1413;         | Conservative  | 0; Mismatches 192;  | Indels 3; Gaps 2;          |
| Qy                    | 85  | GTCAATATTAGAACGAAGTCCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG       | 144                        |
| Db                    | 250   | GTCACTATTTAGAACGAAGATGCTTGAATGAGAGAGATTATCGGACCGGAGATACATT      | 309                        |
| Qy                    | 145   | ACGAATACGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCAGAGACATTG        | 204                        |
| Db                    | 310   | ATGAATACGAATGACTACTGCGAAGGATATGTTTCCAGACATTACCATAGAGCGTTG       | 369                        |
| Qy                    | 205   | AAAGCGGGTATCGAATCCACTGCAAGTAAATCTTTTCAGTCCGACGAGAGAACGAGTCCTA   | 264                        |
| Db                    | 370   | AAAGCACTTACCGGATCCATTGTCAGTAAATCTCTCAGTCAGGAGCAGGAGACGCCCTA     | 429                        |
| Qy                    | 265   | AAAGAACCGCATAGACACTGTTCAAGTCATCAGTCACGTTCCGAGAGCCACCGAGGA       | 324                        |
| Db                    | 430   | AGAGAAAGCGTAAATAGACCCCTGTGCAAGTATCATCAGTCGCATTGAAGAGCCACCGAAGGA | 489                        |
| Qy                    | 325   | AAAGATCCAGGAGTATAGAGATGATCAGAGGGGTACCTGATCTGTCAAAGTGGAGACG      | 384                        |
| Db                    | 490   | AAAGATCCAGGAGTATAGAGATGATCAGAGGGGTACCTGATCTGTCAAAGTGGAGACG      | 549                        |
| Qy                    | 385   | TTCTTAAGAGCAAGATATGAATCGTGAACACTTTTGGGTGAAGAGCCCTTTGGCAAGTTG    | 444                        |







**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 01:13:38 ; Search time 1244.4 Seconds  
(without alignments)  
11198.198 Million cell updates/sec

Title: US-10-801-671-1

Perfect score: 2354

Sequence: 1 gccagcgggtactttaa.....taaaaaaaaaaaaaaac 2354

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*  
1: geneseqn1980s.\*  
2: geneseqn1990s.\*  
3: geneseqn2000s.\*  
4: geneseqn2001as.\*  
5: geneseqn2001bs.\*  
6: geneseqn2002as.\*  
7: geneseqn2002bs.\*  
8: geneseqn2003as.\*  
9: geneseqn2003bs.\*  
10: geneseqn2003cs.\*  
11: geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 2354   | 100.0       | 2354   | 6     | Aad32038 Human kin |
| 2          | 2215.8 | 94.1        | 2497   | 4     | Aaf89402 Human cel |
| 3          | 2052.6 | 87.2        | 2446   | 2     | Aax89852 Human pro |
| 4          | 2048   | 87.0        | 2488   | 4     | Aaf44675 Novel pro |
| 5          | 2048   | 87.0        | 2488   | 12    | Adi29373 Human MAR |
| 6          | 1632   | 69.3        | 1814   | 13    | Adp24393 PRO polyp |
| 7          | 1629.4 | 69.2        | 1792   | 12    | Adi26142 Human CDN |
| 8          | 1300.6 | 55.3        | 1323   | 3     | Aac81772 Human bea |
| 9          | 1286   | 54.6        | 1446   | 12    | Ado07809 Human pol |
| 10         | 1277.4 | 54.3        | 1881   | 12    | Adi26140 Human CDN |
| 11         | 1185   | 50.3        | 1222   | 3     | Aac59283 Human sec |
| 12         | 1178.2 | 50.1        | 1456   | 11    | Adi31373 Human CDN |
| 13         | 1118.8 | 47.5        | 1865   | 12    | Adi26144 Human CDN |
| 14         | 1097.2 | 46.6        | 1446   | 12    | Ado07812 Human pol |
| 15         | 1097   | 46.6        | 21234  | 6     | Aad32039 Human kin |
| 16         | 813.2  | 34.5        | 1455   | 12    | Ado07810 Human pol |
| 17         | 811.6  | 34.5        | 1834   | 12    | Adh58708 Human CDC |
| 18         | 811.6  | 34.5        | 1834   | 13    | Adri4000 Human NF- |
| 19         | 811.6  | 34.5        | 1834   | 13    | Adp24718 PRO polyp |
| 20         | 768.8  | 32.7        | 1452   | 12    | Ado07813 Mouse pol |

|    |       |      |      |    |                    |
|----|-------|------|------|----|--------------------|
| 21 | 768.4 | 32.6 | 906  | 12 | Ado07814 Mouse pol |
| 22 | 669.2 | 28.4 | 1785 | 10 | Adg74696 Human kin |
| 23 | 644.4 | 27.4 | 1743 | 12 | Adh58715 Human CDC |
| 24 | 643.2 | 27.3 | 1750 | 6  | Abk84044 Human CDN |
| 25 | 643.2 | 27.3 | 1750 | 10 | Adl24750 Intestina |
| 26 | 640.6 | 27.2 | 2254 | 12 | Adg77073 Human CDN |
| 27 | 640   | 27.2 | 2516 | 3  | Aaf21734 Human bre |
| 28 | 630.6 | 26.8 | 1643 | 10 | Adl24751 Intestina |
| 29 | 499.6 | 21.2 | 1500 | 10 | Acag3031 Human clk |
| 30 | 499.6 | 21.2 | 1973 | 6  | Abi64388 Stomach c |
| 31 | 499.6 | 21.2 | 1973 | 6  | Abn95692 Gene #219 |
| 32 | 499.6 | 21.2 | 1973 | 12 | Adq15173 Human can |
| 33 | 499.6 | 21.2 | 1973 | 12 | Adq83230 Human tum |
| 34 | 499.6 | 21.2 | 1973 | 12 | Adq83910 Human tum |
| 35 | 499.6 | 21.2 | 1973 | 12 | Adq84393 Human tum |
| 36 | 499.6 | 21.2 | 1973 | 12 | Adq85141 Human tum |
| 37 | 499.6 | 21.2 | 1973 | 12 | Adq86323 Human tum |
| 38 | 499.6 | 21.2 | 2111 | 9  | Aad57339 Human kin |
| 39 | 464.4 | 19.7 | 1296 | 4  | Aah46906 CDNA enco |
| 40 | 464.4 | 19.7 | 1842 | 5  | Aas87541 DNA enco  |
| 41 | 461.6 | 19.6 | 1026 | 12 | Ado07807 Human pol |
| 42 | 458.6 | 19.5 | 1885 | 12 | Adj62819 Human CDN |
| 43 | 452.2 | 19.2 | 1473 | 12 | Ado07811 Mouse pol |
| 44 | 443   | 18.8 | 443  | 3  | Aac81771 Human bea |
| 45 | 441.4 | 18.8 | 2905 | 8  | Acc46217 Human dit |

ALIGNMENTS

RESULT 1  
AAD32038  
ID AAD32038 standard; cDNA; 2354 BP.  
XX  
AC AAD32038;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human kinase protein CDNA.  
XX  
KW Human; kinase protein; enzyme; cytotatic; osteopathic; gene expression;  
KW colon-moderately differentiated adenocarcinoma; chromosome mapping; gene;  
KW drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;  
KW bone osteosarcoma; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 5'UTR 1..32 /\*tag= a  
FT CDS 33..1370 /\*tag= b  
FT FT /\*product= "Human kinase protein"  
FT 3'UTR 1371..2354 /\*tag= c  
FT WO200216567-A2.  
PD 28-FEB-2002.  
XX  
PF 24-AUG-2001; 2001WO-US026389.  
XX  
PR 24-AUG-2000; 2000US-0227470P.  
PR 19-MAR-2001; 2001US-00810671.  
XX  
PA (APPL-) APPLERA CORP.  
PI Yan C, Ye J, Ketchum KA, Di Francesco V, Beasley EM;  
XX  
DR WPI; 2002-269354/31.  
DR P-PSDB; AAE20170.  
XX  
PT New human kinase proteins and nucleic acids, useful in drug screening

assays, identifying modulators of kinase activity or treating disorders characterized by absence or unwanted expression of the protein.  
 Claim 4; Fig 1; 81pp; English.

The invention relates to isolated human kinase proteins and nucleic acids. The nucleic acid and peptide sequences can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate kinase activity in cells and tissues that express the kinase. The nucleic acids are useful as probes or primers, in constructing recombinant vectors, for expressing antigenic portions of the proteins, chromosome mapping, drug screening, testing an individual for a genotype, and for gene therapy in patients containing cells that are aberrant in kinase gene expression. The proteins may be used in drug screening assays, in the identification of compounds that modulate, stimulate or inhibit kinase activity, in pharmacogenomic analysis, in treating disorders characterized by an absence or unwanted expression of the protein (bone osteosarcoma, or colon-moderately differentiated adenocarcinoma), and in generating antibodies specific for the peptides. Such antibodies can be used to detect the protein in situ, in vitro, or in cell lysate or supernatant, to isolate and purify the proteins from host cells, pharmacogenomic analysis, tissue typing, and in inhibiting protein function. The present sequence is human protein cDNA. Human kinase protein gene is located on chromosome 5

Sequence 2354 BP; 783 A; 370 C; 472 G; 729 T; 0 U; 0 Other;

Query Match 100.0%; Score 2354; DB 6; Length 2354;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | GCCAGCTGGGTACTTTAAACACATGCTCCATGTGCATCCCTTGAAGCTTCGCACT      | 60  |
| DB | 1   | GCCAGCTGGGTACTTTAAACACATGCTCCATGTGCATCCCTTGAAGCTTCGCACT      | 60  |
| QY | 61  | CTGTTGAAGAGGACACTCATCCAGTCAATATTTAGAACGAGGTCCTTGAATGAGCGAG   | 120 |
| DB | 61  | CTGTTGAAGAGGACACTCATCCAGTCAATATTTAGAACGAGGTCCTTGAATGAGCGAG   | 120 |
| QY | 121 | ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAGGATATGTTTC  | 180 |
| DB | 121 | ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAGGATATGTTTC  | 180 |
| QY | 181 | CTAGACATTTACACAGACACATTGAAAGCGGTATCGAATCCACTGACAGTAAATCTTCAG | 240 |
| DB | 181 | CTAGACATTTACACAGACACATTGAAAGCGGTATCGAATCCACTGACAGTAAATCTTCAG | 240 |
| QY | 241 | TCCGACGAGGAGGACAGTCTCTAAAGGAGCGCAATAGACACTGTTTCAAGTCATCAGT   | 300 |
| DB | 241 | TCCGACGAGGAGGACAGTCTCTAAAGGAGCGCAATAGACACTGTTTCAAGTCATCAGT   | 300 |
| QY | 301 | CACGTTTGAAGAGGACCGAAGGAAAGATCAGGAGTATAGAGGATGATGAGGAGGTC     | 360 |
| DB | 301 | CACGTTTGAAGAGGACCGAAGGAAAGATCAGGAGTATAGAGGATGATGAGGAGGTC     | 360 |
| QY | 361 | ACCTGATCTGTCAGAGTGGAGACGTTCTTAAGAGCAAGATATGAATCTGTGACACTTTGG | 420 |
| DB | 361 | ACCTGATCTGTCAGAGTGGAGACGTTCTTAAGAGCAAGATATGAATCTGTGACACTTTGG | 420 |
| QY | 421 | GTGAGGAGGCTTTGGCAAGTGTAGAGTGCATTTGATCATGGCATGGATGGCATGCATG   | 480 |
| DB | 421 | GTGAGGAGGCTTTGGCAAGTGTAGAGTGCATTTGATCATGGCATGGATGGCATGCATG   | 480 |
| QY | 481 | TAGCAGTGAATCGTAAACATGTAGCCGTTACCGTGAAGCAGCTCGTTCAGAAATCC     | 540 |
| DB | 481 | TAGCAGTGAATCGTAAACATGTAGCCGTTACCGTGAAGCAGCTCGTTCAGAAATCC     | 540 |
| QY | 541 | AAGTATTAGACACCTTAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGC  | 600 |
| DB | 541 | AAGTATTAGACACCTTAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGC  | 600 |
| QY | 601 | TAGAATGTTTGCATCATGTGTCATGTTTGTATGTTTGAACACTACGCGACTTAGTA     | 660 |

|    |      |  |      |
|----|------|--|------|
| DB | 601  | TAGAATGTTTGCATCATGTGTCATGTTTGTATGTTTGAACACTACGCGACTTAGTA       | 660  |
| QY | 661  | CTTACGATTTTCAATTAAGAAACAGCTTTCTGCCATTTCAAAATGACACACATCAGGAGA   | 720  |
| DB | 661  | CTTACGATTTTCAATTAAGAAACAGCTTTCTGCCATTTCAAAATGACACACATCAGGAGA   | 720  |
| QY | 721  | TGGGATATCAGATCGCCAGTCAATAAATTTTACATCATATAAATAAATTAACCCATACAG   | 780  |
| DB | 721  | TGGGATATCAGATCGCCAGTCAATAAATTTTACATCATATAAATAAATTAACCCATACAG   | 780  |
| QY | 781  | ATCTGAAGCCTGAAATATTTTGTGTAAGTCTGACTATGTAGTCAAAATATTAATCTTA     | 840  |
| DB | 781  | ATCTGAAGCCTGAAATATTTTGTGTAAGTCTGACTATGTAGTCAAAATATTAATCTTA     | 840  |
| QY | 841  | AAATGAACCTGATGAACGACACTGAAACACAGATATCAAAAGTTGTGTGACCTTGGAA     | 900  |
| DB | 841  | AAATGAACCTGATGAACGACACTGAAACACAGATATCAAAAGTTGTGTGACCTTGGAA     | 900  |
| QY | 901  | GTGCAACGTATGATGAACACATCAACAGTACTTTGGTGTCTACCCGGCCTACACAGCTC    | 960  |
| DB | 901  | GTGCAACGTATGATGAACACATCAACAGTACTTTGGTGTCTACCCGGCCTACACAGCTC    | 960  |
| QY | 961  | CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTGGAGCATAGTTGCA     | 1020 |
| DB | 961  | CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTGGAGCATAGTTGCA     | 1020 |
| QY | 1021 | TTCTTATTTGAATATTTACCTTGGTTTTCAGAGTCTTTTACAGTCTCATGATAGTAAGACAC | 1080 |
| DB | 1021 | TTCTTATTTGAATATTTACCTTGGTTTTCAGAGTCTTTTACAGTCTCATGATAGTAAGACAC | 1080 |
| QY | 1081 | TGGCAATGATGGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAAAAGAA     | 1140 |
| DB | 1081 | TGGCAATGATGGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAAAAGAA     | 1140 |
| QY | 1141 | AACCAAGTATTTTACCATTAACAGCTAGATTGGGATGACACAGTCTTCTGCTGGTAGAT    | 1200 |
| DB | 1141 | AACCAAGTATTTTACCATTAACAGCTAGATTGGGATGACACAGTCTTCTGCTGGTAGAT    | 1200 |
| QY | 1201 | ATGTTTAGAGACGCTGCAACCGTTGAAGGAATTTATGCTTGTTCATGATCAAGAAACATG   | 1260 |
| DB | 1201 | ATGTTTAGAGACGCTGCAACCGTTGAAGGAATTTATGCTTGTTCATGATCAAGAAACATG   | 1260 |
| QY | 1261 | AGAACTGTTTGCACCTGTTGGAAGATTTTGTAGAAATATGATCCAACTCAAAAGATTAAC   | 1320 |
| DB | 1261 | AGAACTGTTTGCACCTGTTGGAAGATTTTGTAGAAATATGATCCAACTCAAAAGATTAAC   | 1320 |
| QY | 1321 | TGGATGAAGCATTTGACGATCCTTTTTCGAAAGATTTAGAAATATGATCCAACTCAAAAG   | 1380 |
| DB | 1321 | TGGATGAAGCATTTGACGATCCTTTTTCGAAAGATTTAGAAATATGATCCAACTCAAAAG   | 1380 |
| QY | 1381 | AGTGGTCTTACTATATCTCTTAGAAGAGATTTTAAAGTCTGTGTGATCAACTAAA        | 1440 |
| DB | 1381 | AGTGGTCTTACTATATCTCTTAGAAGAGATTTTAAAGTCTGTGTGATCAACTAAA        | 1440 |
| QY | 1441 | CATTCTAATATTTTGTAAACATTTAAATTTTGTACAGTTAAGTGTAAATATTTGTATG     | 1500 |
| DB | 1441 | CATTCTAATATTTTGTAAACATTTAAATTTTGTACAGTTAAGTGTAAATATTTGTATG     | 1500 |
| QY | 1501 | TTTTTGTATCAATAGATAAATTAACCTTGTAGCAAGTATGCTTGTGATTAAGTCAATGAA   | 1560 |
| DB | 1501 | TTTTTGTATCAATAGATAAATTAACCTTGTAGCAAGTATGCTTGTGATTAAGTCAATGAA   | 1560 |
| QY | 1561 | AAATTAATAATTTTCTTTTGAATTAACCTTTTAAATACCTTTTGAATATCTCTT         | 1620 |
| DB | 1561 | AAATTAATAATTTTCTTTTGAATTAACCTTTTAAATACCTTTTGAATATCTCTT         | 1620 |
| QY | 1621 | TGTGTCCAGTGAATATGATGATCTTGTGCTTTGTGATGATGAGGTCACCTCTGAAGT      | 1680 |
| DB | 1621 | TGTGTCCAGTGAATATGATGATCTTGTGCTTTGTGATGATGAGGTCACCTCTGAAGT      | 1680 |
| QY | 1681 | GATTTTCTTTCAGTAAAGGAAATCTTTCAGTCTTTTATTTTCTTAAAGCAATATCTTTTA   | 1740 |



Db 1681 GATTTTTTTGAGTAAAGGAAATCTTGACTACTATTTATTTCTTTAAAGGAATATTTCTTTA 1740  
Qy 1741 TATACCTTCAAAATTTAGAACTTTAACTTTTAAAGTTTCTTCTGTAATTTGTAACGGGTG 1800  
Db 1741 TATACCTTCAAAATTTAGAACTTTAACTTTTAAAGTTTCTTCTGTAATTTGTAACGGGTG 1800  
Qy 1801 ATTATTTAATCTCTAGATAGCAGGTACTAGAAACCAAACTCAGAAATGTTTACTGT 1860  
Db 1801 ATTATTTAATCTCTAGATAGCAGGTACTAGAAACCAAACTCAGAAATGTTTACTGT 1860  
Qy 1861 TAGAATCTTATTAATTTTAAAGTTTAAAGTTTCTTCTGTAATTTGTAACGGGTGATA 1920  
Db 1861 TAGAATCTTATTAATTTTAAAGTTTAAAGTTTCTTCTGTAATTTGTAACGGGTGATA 1920  
Qy 1921 ACCAGACATTCATGAAAGGATGAGTTTGTCCATTTGTGACAGTTTCTTTTAAATAAAC 1980  
Db 1921 ACCAGACATTCATGAAAGGATGAGTTTGTCCATTTGTGACAGTTTCTTTTAAATAAAC 1980  
Qy 1981 ACATACACATCTTATTTAAGATTTAAATCTAACTGGAAGTCAGCTTGGAAATGGACAT 2040  
Db 1981 ACATACACATCTTATTTAAGATTTAAATCTAACTGGAAGTCAGCTTGGAAATGGACAT 2040  
Qy 2041 TTCCAAGTATCTTGTGAGTCAAGATTAATAAATAGAAATTTCTGATGAGAGTTTCAG 2100  
Db 2041 TTCCAAGTATCTTGTGAGTCAAGATTAATAAATAGAAATTTCTGATGAGAGTTTCAG 2100  
Qy 2101 TTTTAAATACCAAGTCCTTAGGAGTCTTAACTGAGGAGATCTGTTTATCAAAATGACA 2160  
Db 2101 TTTTAAATACCAAGTCCTTAGGAGTCTTAACTGAGGAGATCTGTTTATCAAAATGACA 2160  
Qy 2161 TAAATACGTAACCTTAAAGATTTAAAGTTTAAATTTAATTTAGGCAATTTATGCTGTGATAAT 2220  
Db 2161 TAAATACGTAACCTTAAAGATTTAAAGTTTAAATTTAATTTAGGCAATTTATGCTGTGATAAT 2220  
Qy 2221 TCTTACGGGAGAAAGAGATTTGATTTGAAAGCAGTTTGGGAGAAAGTCTGCTGAAAT 2280  
Db 2221 TCTTACGGGAGAAAGAGATTTGATTTGAAAGCAGTTTGGGAGAAAGTCTGCTGAAAT 2280  
Qy 2281 TTCCAGAAATTTAATTTGATTTGATTTGAAAGCAGTTTGGGAGAAAGTCTGCTGAAAT 2340  
Db 2281 TTCCAGAAATTTAATTTGATTTGATTTGAAAGCAGTTTGGGAGAAAGTCTGCTGAAAT 2340  
Qy 2341 AACCAAAAAAACC 2354  
Db 2341 AACCAAAAAAACC 2354

RESULT 2  
AAF89402  
ID AAF89402 standard; cDNA; 2497 BP.  
XX AC AAF89402;  
XX AC AAF89402;  
XX DT 14-AUG-2001 (first entry)  
XX DE Human cell cycle regulating protein 53 coding sequence.  
XX DE Human; cell cycle regulating protein 53; cancer; blood disease; HIV;  
XX KW immunological disease; inflammation; ss.  
XX OS Homo sapiens.  
XX OS WO200130833-A1.  
XX PN 03-MAY-2001.  
XX PD 16-OCT-2000; 2000WO-CN000328.  
XX PF 22-OCT-1999; 99CN-00119816.  
XX PR (SHAN-) SHANGHAI BIO ROAD GENE DEV LTD.  
XX PA Mao Y, Xie Y;  
XX PI

XX WPI; 2001-300480/31.  
DR P-PSDB; AAB69705.  
XX New cell cycle-regulating protein 53 and its polynucleotide, applicable  
PT in diagnosis and treatment of malignant tumor; hemopathy, human  
PT immunodeficiency virus infection, immunological diseases and various  
XX inflammation.  
XX Claim 6; Page 22-23; 30pp; Chinese.  
XX The present invention provides the protein and coding sequences for human  
CC cell cycle regulating protein 53. The sequences can be used in the  
CC diagnosis and treatment of malignant tumours, haemopathy, human  
CC immunodeficiency virus (HIV) infection, immunological diseases and  
CC various types of inflammation. The present sequence is the cell cycle  
CC regulating protein 53 coding sequence  
XX Sequence 2497 BP; 801 A; 399 C; 542 G; 755 T; 0 U; 0 Other;  
SQ Query Match 94.1%; Score 2215.8; DB 4; Length 2497;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2220; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 85 GTCATTTATTAGAACGAGGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 144  
Db 271 GTCATTTATTAGAACGAGGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 330  
Qy 145 ACGAATACGGAATGACTACTGTCGAGGATATGTTCTTAGACATTATCAGAGACATTG 204  
Db 331 ACGAATACGGAATGACTACTGTCGAGGATATGTTCTTAGACATTATCAGAGACATTG 390  
Qy 205 AAAGCGGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGAGCAGGAGAGAGTCTCTA 264  
Db 391 AAAGCGGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGAGCAGGAGAGAGTCTCTA 450  
Qy 265 AAAGGAGCGCAATAGACATCTGTTCAAGTCACTAGTCAGTTCGAGAGCCACCGAAGGA 324  
Db 451 AAAGGAGCGCAATAGACATCTGTTCAAGTCACTAGTCAGTTCGAGAGCCACCGAAGGA 510  
Qy 325 AAAGTCCAGGATGATGAGGATGATGAGGAGGTGACCTGATCTGTCAAAAGTGAGAGC 384  
Db 511 AAAGTCCAGGATGATGAGGATGATGAGGAGGTGACCTGATCTGTCAAAAGTGAGAGC 570  
Qy 385 TTCTAAGAGCAAGATATGAATCTGTGACACCTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 444  
Db 571 TTCTAAGAGCAAGATATGAATCTGTGACACCTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 630  
Qy 445 TAGAGTGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 504  
Db 631 TAGAGTGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690  
Qy 505 TAGGCGGTACCGTGAAGCAGCTCGTTTCAGAAATCCAGATTTAGAGCAGTAAATAGTA 564  
Db 691 TAGGCGGTACCGTGAAGCAGCTCGTTTCAGAAATCCAGATTTAGAGCAGTAAATAGTA 750  
Qy 565 CTGATCCCAATAGTGTCTTCCGATGTCGTCAGATGCTAGAAATGTTGATCATCATGTC 624  
Db 751 CTGATCCCAATAGTGTCTTCCGATGTCGTCAGATGCTAGAAATGTTGATCATCATGTC 810  
Qy 625 ATGTTTGTATTTGTTGAACTACTGGAAGTCTAGTACTTACGATTTTCAATTAAGAAACA 684  
Db 811 ATGTTTGTATTTGTTGAACTACTGGAAGTCTAGTACTTACGATTTTCAATTAAGAAACA 870  
Qy 685 GCTTCTGCCCATTTCAATTTGACCATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 744  
Db 871 GCTTCTGCCCATTTCAATTTGACCATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 930  
Qy 745 TAAATTTTATACATCAATAAATAAATTAACCATCAGATCTGAAGCCTGAAATATTTTGT 804  
Db 931 TAAATTTTATACATCAATAAATAAATTAACCATCAGATCTGAAGCCTGAAATATTTTGT 990  
Qy 805 TTGTGAAGTCTGACTATGTAGTCAATAATATTTCTAAATTAAGAAAGTGAACGACAC 864

Db 991 TTGTGAAGTCTGACATATGTAGTCAAAATATAATCTTAAATGAAACGTGATGAAGCCAC 1050  
Qy 865 TGAATAACACAGATATCAAAAGTGTGTGACTTTGGAGTGAACAGTATGATGATGAATC 924  
Db 1051 TGAATAACACAGATATCAAAAGTGTGTGACTTTGGAGTGAACAGTATGATGATGAATC 1110  
Qy 925 ACAGTACTTTTGGTGTCTACCGGCACTACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT 984  
Db 1111 ACAGTACTTTTGGTGTCTACCGGCACTACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT 1170  
Qy 985 GGTCTCAGCCTTTGGATGTTTGGAGCATAGGTGCAATCTTATGAAATATTACCTTGGTT 1044  
Db 1171 GGTCTCAGCCTTTGGATGTTTGGAGCATAGGTGCAATCTTATGAAATATTACCTTGGTT 1230  
Qy 1045 TCACAGTCTTCAGACTCATGATAGTAAAGACCTCGGCAATGATGGAACGAATATTAG 1104  
Db 1231 TCACAGTCTTCAGACTCATGATAGTAAAGACCTCGGCAATGATGGAACGAATATTAG 1290  
Qy 1105 GACCCATACCAACACACATGATTCAGAAAAACAAGAAAAACGAAGTATTTCACCATTAACC 1164  
Db 1291 GACCCATACCAACACACATGATTCAGAAAAACAAGAAAAACGAAGTATTTCACCATTAACC 1350  
Qy 1165 AGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1224  
Db 1351 AGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1410  
Qy 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAAACTGTTGACCTGGTTCGAA 1284  
Db 1411 TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAAACTGTTGACCTGGTTCGAA 1470  
Qy 1285 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTCGATGAAGCAATTCGAGATCCCT 1344  
Db 1471 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTCGATGAAGCAATTCGAGATCCCT 1530  
Qy 1345 TCTTTGACTTATTAATAAGAAATGAAATGGAATCAGTGGTCTTACTATATACCTTCTCT 1404  
Db 1531 TCTTTGGCTTATTAATAAGAAATGAAATGGAATCAGTGGTCTTACTATATACCTTCTCT 1590  
Qy 1405 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTCAATCAATCTTAATATTTTGTAAACAT 1464  
Db 1591 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTCAATCAATCTTAATATTTTGTAAACAT 1650  
Qy 1465 AAATTAATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATTAATAAC 1524  
Db 1651 AAATTAATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATTAATAAC 1710  
Qy 1525 TTGTTAAGCAAGTATGGTCTTTGATATGCAATAGAAAAATTAATAATTTTCTTTT 1584  
Db 1711 TTGTTAAGCAAGTATGGTCTTTGATATGCAATAGAAAAATTAATAATTTTCTTTT 1770  
Qy 1585 GAAATTACCAATTTTAATACCTTTGAATATCCTTTGTGTCAGTGAATAATGTGATTG 1644  
Db 1771 GAAATTACCAATTTTAATACCTTTGAATATCCTTTGTGTCAGTGAATAATGTGATTG 1830  
Qy 1645 ATCTTGCTTTTGTACATGGAGGTCACTCTGAGTGAATTTTGTAGTAAAGGAAT 1704  
Db 1831 ATCTTGCTTTTGTACATGGAGGTCACTCTGAGTGAATTTTGTAGTAAAGGAAT 1890  
Qy 1705 CTTGACTACTTTATATCTTAAAGGAATTTCTTTATATACCTCAAAATTTAGAACTTAAC 1764  
Db 1891 CTTGACTACTTTATATCTTAAAGGAATTTCTTTATATACCTCAAAATTTAGAACTTAAC 1950  
Qy 1765 TTTAAAGTTTTTCTCTGATATTTGTAAGGGGTATATATTAATCTAGATAAGCA 1824  
Db 1951 TTTAAAGTTTTTCTCTGATATTTGTAAGGGGTATATATTAATCTAGATAAGCA 2010  
Qy 1825 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGT 1884  
Db 2011 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGT 2070  
Qy 1885 TTGTATTTCTTTTCAATTCGGTGTGTCAGGGTGAATAACGACATTCATGGAAGGCGATG 1944

Db 2071 TTGTATTTCTTTTTCATTTGGGTGATGTCAGGGTGATTAACGACACATTCATGGAAGGCGATG 2130  
Qy 1945 CAGTTTGTCCATTTGTGACAGTTTCTTTTAATAAAACCATACACACTTTTATTTAAGATTA 2004  
Db 2131 CAGTTTGTCCATTTGTGACAGTTTCTTTTAATAAAACCATACACACTTTTATTTAAGATTA 2190  
Qy 2005 AAATCTAACTCGAAAGTCAGCTTCGAAAAATGCAATTTCCAAAGTATGTTTGGTGAAGTCAC 2064  
Db 2191 AAATCTAACTCGAAAGTCAGCTTCGAAAAATGCAATTTCCAAAGTATGTTTGGTGAAGTCAC 2250  
Qy 2065 AGATATAAAATAGAAATTCGTATGAGAGGTTTCAGTTTTTAATACCAAGTCCCTTAGAG 2124  
Db 2251 AGATATAAAATAGAAATTCGTATGAGAGGTTTCAGTTTTTAATACCAAGTCCCTTAGAG 2310  
Qy 2125 TCTTAAATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTTATAGAAT 2184  
Db 2311 TCTTAAATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTTATAGAAT 2370  
Qy 2185 AAGTTTATTAATAGGCAATTTATGTCGTGATAATCTTACGGGAGAAAGAGGATTTGA 2244  
Db 2371 AAGTTTATTAATAGGCAATTTATGTCGTGATAATCTTACGGGAGAAAGAGGATTTGA 2430  
Qy 2245 TTGAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAATGATTTGGTTA 2304  
Db 2431 TTGAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAATGATTTGGTTA 2490  
Qy 2305 CATAAAC 2311  
Db 2491 CATAAAC 2497

RESULT 3  
AAx89852  
ID AAx89852 standard; cdna; 2446 BP.  
XX  
AC AAx89852;  
XX  
DT 08-OCT-1999 (first entry)  
XX  
DE Human protein kinase (HPKM)-3 encoding nucleotide (clone ID 339963).  
XX  
KW Human protein kinase molecule; HPKM; human; protein kinase;  
phosphate group; cancer; immune disorder; ss.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 128..1483  
FT /\*tag= a  
FT /product= "HPKM-3"  
XX  
PN WO9938981-A2.  
XX  
PD 05-AUG-1999.  
XX  
PF 12-JAN-1999; 99WO-US000661.  
XX  
PR 30-JAN-1998; 98US-00016000.  
XX  
PA (INCY-) INCYTE PHARM INC.  
XX  
PI Bandman O, Hillman JL, Lal P, Akerblom IE, Shah P, Corley NC;  
PI Guegler KJ;  
XX  
XX WPI; 1999-479190/40.  
DR P-ESDB; AAY27054.  
XX  
PT New human protein kinase molecules useful for treating or preventing  
cancer or an immune disorder.  
PS Claim 8; Page 74-75; 77pp; English.  
XX  
CC The invention provides human protein kinase molecules (HPKM) (AAY27052-

|    |   |
|----|---|
| CC | 57) and nucleic acid sequences (AA89850-55) encoding the HPKM             |
| CC | polypeptides respectively. The HPKM polypeptides can be produced          |
| CC | recombinantly by standard recombinant methodology. Protein kinases add    |
| CC | phosphate groups to proteins. HPKM polypeptide or an HPKM antagonist is   |
| CC | used to treat or prevent cancer or an immune disorder. The present        |
| CC | sequence represents the nucleotide sequence of HPKM-3                     |
| XX |   |
| SQ | Sequence 2446 BP; 788 A; 391 C; 516 G; 751 T; 0 U; 0 Other;               |
|    | Query Match 87.2%; Score 2052.6; DB 2; Length 2446;                       |
|    | Best Local Similarity 95.8%; Pred. No. 0;                                 |
|    | Matches 2155; Conservative 0; Mismatches 4; Indels 90; Gaps 1;            |
| QY | 85 GTCAATATTTAGAACCAAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 144        |
| DB | 288 GTCAATATTTAGAACCAAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 347       |
| QY | 145 ACGAATACAGGATGACCTACTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG 204      |
| DB | 348 ACGAATACAGGATGACCTACTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG 407      |
| QY | 205 AAAGCGGTATCGAATCGAATCGAATCTTCAAGTCCGACGAGGAGAGCAGTCCCTA 264           |
| DB | 408 AAAGCGGTATCGAATCGAATCGAATCTTCAAGTCCGACGAGGAGAGCAGTCCCTA 467           |
| QY | 265 AAAGGAAGCGCAATAGACACTGTTTCAAGTCACTCAGTTCAGTTCGAAAGGACCCGAAGGA 324     |
| DB | 468 AAAGGAAGCGCAATAGACACTGTTTCAAGTCACTCAGTTCGAAAGGACCCGAAGGA 515          |
| QY | 325 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGCG 384       |
| DB | 516 ----- 515   |
| QY | 385 TTCTAAGAGCAAGATATGAATCTGGACACTTTGGGTGAGGAGCCTTTGCGCAAGTTG 444         |
| DB | 516 -----AATCCGTGGACACTTTGGGTGAGGAGCCTTTGCGCAAGTTG 557                    |
| QY | 445 TAGAGTGCATTTGATCGGATGAGTGGATGATGATGATGATGATGATGATGATGATGATGATG 504    |
| DB | 558 TAGAGTGCATTTGATCGGATGAGTGGATGATGATGATGATGATGATGATGATGATGATGATG 617    |
| QY | 505 TAGGCGGTACCGTGAAGCAGCTCTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 564     |
| DB | 618 TAGGCGGTACCGTGAAGCAGCTCTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 677     |
| QY | 565 CTGATCCCAATAGTCTCTTCGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 624   |
| DB | 678 CTGATCCCAATAGTCTCTTCGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 737   |
| QY | 625 ATGTTTGTATTGTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 684        |
| DB | 738 ATGTTTGTATTGTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 797        |
| QY | 685 GCTTTCTGCCATTTCAAAATGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 744        |
| DB | 798 GCTTTCTGCCATTTCAAAATGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 857        |
| QY | 745 TAAATTTTACATCATTAATAATTAACCCATACAGATCTGAGGCTGAAAATATTTTGT 804         |
| DB | 858 TAAATTTTACATCATTAATAATTAACCCATACAGATCTGAGGCTGAAAATATTTTGT 917         |
| QY | 805 TTGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 864   |
| DB | 918 TTGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 977   |
| QY | 865 TGAATAACACAGATATCAAAAGTTGTTGACTTTTGAAGTGCACCGTATGATGATGATGATGATG 924  |
| DB | 978 TGAATAACACAGATATCAAAAGTTGTTGACTTTTGAAGTGCACCGTATGATGATGATGATGATG 1037 |
| QY | 925 ACAGTACTTTGGTGTCTACCCGGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 984    |
| DB | 1038 ACAGTACTTTGGTGTCTACCCGGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 1097  |

|    |  |
|----|--|
| QY | 985 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGCATTTCTTATTGAATATTACCTTGGTT 1044       |
| DB | 1098 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGCATTTCTTATTGAATATTACCTTGGTT 1157      |
| QY | 1045 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCAGCTGGCATGATGAGCAAGATATTAG 1104        |
| DB | 1158 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCAGCTGGCATGATGAGCAAGATATTAG 1217        |
| QY | 1105 GACCCATACCAACACACATGATTTCAGAAAAACAGAAAAACGCAAGTATTTTTCCACATAACC 1164    |
| DB | 1218 GACCCATACCAACACACATGATTTCAGAAAAACAGAAAAACGCAAGTATTTTTCCACATAACC 1277    |
| QY | 1165 AGCTAGATTGGGATGAACACACAGTCTTCTGCTGTAGATATGTTAGGAGAGCGTGCAACCGT 1224     |
| DB | 1278 AGCTAGATTGGGATGAACACACAGTCTTCTGCTGTAGATATGTTAGGAGAGCGTGCAACCGT 1337     |
| QY | 1225 TGAAGGAATTTATGCTTTGTCTATGATGAAGAACATGAGAAACTGTTTTCACCTGGTTCGAA 1284     |
| DB | 1338 TGAAGGAATTTATGCTTTGTCTATGATGAAGAACATGAGAAACTGTTTTCACCTGGTTCGAA 1397     |
| QY | 1285 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTCGATGAAGCATTTGCAGCATCCTT 1344    |
| DB | 1398 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTCGATGAAGCATTTGCAGCATCCTT 1457    |
| QY | 1345 TCTTTGACTTATTAATAAGAAATGAAATGGGAATCAGTGTCTTATATATCTTCTCT 1404           |
| DB | 1458 TCTTTGACTTATTAATAAGAAATGAAATGGGAATCAGTGTCTTATATATCTTCTCT 1517           |
| QY | 1405 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1464       |
| DB | 1518 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1577       |
| QY | 1465 AAATTTATTTCTGACAGTTTAAAGTAAATATTGATGTTTGTATCAATAGCATATTAAC 1524         |
| DB | 1578 AAATTTATTTCTGACAGTTTAAAGTAAATATTGATGTTTGTATCAATAGCATATTAAC 1637         |
| QY | 1525 TTGTTAAGCAAGTATGGTCTTTGATAATGCAATTAGAAAAATTAATAATTTTCTTTT 1584          |
| DB | 1638 TTGTTAAGCAAGTATGGTCTTTGATAATGCAATTAGAAAAATTAATAATTTTCTTTT 1697          |
| QY | 1585 GAAATACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATATGATG 1644              |
| DB | 1698 GAAATACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATATGATG 1757              |
| QY | 1645 ATCTTGCTTTTGTACATGAGAGTCACCTCTGAAAGTGAATTTTGTAGTAAAGGAAAT 1704          |
| DB | 1758 ATCTTGCTTTTGTACATGAGAGTCACCTCTGAAAGTGAATTTTGTAGTAAAGGAAAT 1817          |
| QY | 1705 CTGACTACTTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAAC 1764     |
| DB | 1818 CTGACTACTTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAAC 1877     |
| QY | 1765 TTTTAAAGTTTTTCTTCTGTAATTTGTTGAACGGGTGATATTTAATCTAGATAAGCA 1824          |
| DB | 1878 TTTTAAAGTTTTTCTTCTGTAATTTGTTGAACGGGTGATATTTAATCTAGATAAGCA 1937          |
| QY | 1825 GGTACTAGAAACCAAACTCAGAAAACTGTTTACGTTAGAAATTTCTTATTAATTTTAAAGTG 1884     |
| DB | 1938 GGTACTAGAAACCAAACTCAGAAAACTGTTTACGTTAGAAATTTCTTATTAATTTTAAAGTG 1997     |
| QY | 1885 TTGTATTTCTTTTCTCAATTTGGGTGATGTCAGGGGTGATAACAGACATTTCTATGAAAGGCATG 1944  |
| DB | 1998 TTGTATTTCTTTTCTCAATTTGGGTGATGTCAGGGGTGATAACAGACATTTCTATGAAAGGCATG 2057  |
| QY | 1945 CAGTTTGTCCATGTCAGAGTTTGTTTAATAAAAAACCATATACACATTTTATTAAAGATTA 2004      |
| DB | 2058 CAGTTTGTCCATGTCAGAGTTTGTTTAATAAAAAACCATATACACATTTTATTAAAGATTA 2117      |
| QY | 2005 AAATCTAACTGGAAGTCAGCTTGGAAAAATGGAATCTTCCAGATGTTTGGTGGATGCAC 2064        |
| DB | 2118 AAATCTAACTGGAAGTCAGCTTGGAAAAATGGAATCTTCCAGATGTTTGGTGGATGCAC 2177        |
| QY | 2065 AGATATAAATAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTTTAAATACCAAGTCTCTTAGGAG 2124 |

|    |      |   |      |
|----|------|---|------|
| Db | 2178 | AGATATAAAAATAGAAATCTGTAGAGAGGTTTCAGTTTTTAATACCAAGTCCTTAGGAG       | 2237 |
| Qy | 2125 | TCCTTAACATTTGGCCAGCATCTGTTTTATCAAAATGACATAAAATACGTAAACCTATAAGAATT | 2184 |
| Db | 2238 | TCCTTAACATTTGGCCAGCATCTGTTTTATCAAAATGACATAAAATACGTAAACCTATAAGAATT | 2297 |
| Qy | 2185 | AAGTTTTATTAATTAGGCAATTTATGCTGTGATAAATCTTACGGCGAGAAAGAGGATTTGA     | 2244 |
| Db | 2298 | AAGTTTTATTAATTAGGCAATTTATGCTGTGATAAATCTTACGGCGAGAAAGAGGATTTGA     | 2357 |
| Qy | 2245 | TTGGAAGACGATTTGGGAAGAAAGTGCCTGCTGAAATTTCCAGAAATTTAAATGATTGGTTA    | 2304 |
| Db | 2358 | TTGGAAGACGATTTGGGAAGAAAGTGCCTGCTGAAATTTCCAGAAATTTAAATGATTGGTTA    | 2417 |
| Qy | 2305 | CATAAACTTTTTTGACTTTCCAGAAAAAAA                                    | 2333 |
| Db | 2418 | CATAAACTTTTTTGACTTTCAAAAAAAA                                      | 2446 |

RESULT 4

RESULI 4  
DAF44675

AAE44675  
ID AAE44675 standard: cDNA: 2488 BP.

XX  
IU  
AAFC44073

AC  
AY

AC  
AA

XX  
DT 27-MAR-2001 (first entry)

DI  
YY

XX DE NOV

DE  
VV  
: NOV

XX  
XX

KW Hum

| KW  | 1mm | 2mm |
|-----|-----|-----|
| 10  | 10  | 10  |
| 20  | 20  | 20  |
| 30  | 30  | 30  |
| 40  | 40  | 40  |
| 50  | 50  | 50  |
| 60  | 60  | 60  |
| 70  | 70  | 70  |
| 80  | 80  | 80  |
| 90  | 90  | 90  |
| 100 | 100 | 100 |

KW  
der

| KW  | imm |
|-----|-----|
| 100 | 100 |
| 90  | 90  |
| 80  | 80  |
| 70  | 70  |
| 60  | 60  |
| 50  | 50  |
| 40  | 40  |
| 30  | 30  |
| 20  | 20  |
| 10  | 10  |
| 0   | 0   |

KW can

KW inf

XX 8

|     |    |
|-----|----|
| Hom | OS |
|-----|----|

XX

PN WO2

XX.

PD 07-

XX

Ed 26-

XX

PR 28-

XX

PA (SU)

22

PI plot

XX  
XX  
XX

| AA | DR | WPI |
|----|----|-----|
|----|----|-----|

DR  
DR  
WET  
P-P

XX  
UR  
F ÷ F

NTC  
PT

| DT  | PI  | NUC | Pre |
|-----|-----|-----|-----|
| 1   | 1   | 1   | 1   |
| 2   | 2   | 2   | 2   |
| 3   | 3   | 3   | 3   |
| 4   | 4   | 4   | 4   |
| 5   | 5   | 5   | 5   |
| 6   | 6   | 6   | 6   |
| 7   | 7   | 7   | 7   |
| 8   | 8   | 8   | 8   |
| 9   | 9   | 9   | 9   |
| 10  | 10  | 10  | 10  |
| 11  | 11  | 11  | 11  |
| 12  | 12  | 12  | 12  |
| 13  | 13  | 13  | 13  |
| 14  | 14  | 14  | 14  |
| 15  | 15  | 15  | 15  |
| 16  | 16  | 16  | 16  |
| 17  | 17  | 17  | 17  |
| 18  | 18  | 18  | 18  |
| 19  | 19  | 19  | 19  |
| 20  | 20  | 20  | 20  |
| 21  | 21  | 21  | 21  |
| 22  | 22  | 22  | 22  |
| 23  | 23  | 23  | 23  |
| 24  | 24  | 24  | 24  |
| 25  | 25  | 25  | 25  |
| 26  | 26  | 26  | 26  |
| 27  | 27  | 27  | 27  |
| 28  | 28  | 28  | 28  |
| 29  | 29  | 29  | 29  |
| 30  | 30  | 30  | 30  |
| 31  | 31  | 31  | 31  |
| 32  | 32  | 32  | 32  |
| 33  | 33  | 33  | 33  |
| 34  | 34  | 34  | 34  |
| 35  | 35  | 35  | 35  |
| 36  | 36  | 36  | 36  |
| 37  | 37  | 37  | 37  |
| 38  | 38  | 38  | 38  |
| 39  | 39  | 39  | 39  |
| 40  | 40  | 40  | 40  |
| 41  | 41  | 41  | 41  |
| 42  | 42  | 42  | 42  |
| 43  | 43  | 43  | 43  |
| 44  | 44  | 44  | 44  |
| 45  | 45  | 45  | 45  |
| 46  | 46  | 46  | 46  |
| 47  | 47  | 47  | 47  |
| 48  | 48  | 48  | 48  |
| 49  | 49  | 49  | 49  |
| 50  | 50  | 50  | 50  |
| 51  | 51  | 51  | 51  |
| 52  | 52  | 52  | 52  |
| 53  | 53  | 53  | 53  |
| 54  | 54  | 54  | 54  |
| 55  | 55  | 55  | 55  |
| 56  | 56  | 56  | 56  |
| 57  | 57  | 57  | 57  |
| 58  | 58  | 58  | 58  |
| 59  | 59  | 59  | 59  |
| 60  | 60  | 60  | 60  |
| 61  | 61  | 61  | 61  |
| 62  | 62  | 62  | 62  |
| 63  | 63  | 63  | 63  |
| 64  | 64  | 64  | 64  |
| 65  | 65  | 65  | 65  |
| 66  | 66  | 66  | 66  |
| 67  | 67  | 67  | 67  |
| 68  | 68  | 68  | 68  |
| 69  | 69  | 69  | 69  |
| 70  | 70  | 70  | 70  |
| 71  | 71  | 71  | 71  |
| 72  | 72  | 72  | 72  |
| 73  | 73  | 73  | 73  |
| 74  | 74  | 74  | 74  |
| 75  | 75  | 75  | 75  |
| 76  | 76  | 76  | 76  |
| 77  | 77  | 77  | 77  |
| 78  | 78  | 78  | 78  |
| 79  | 79  | 79  | 79  |
| 80  | 80  | 80  | 80  |
| 81  | 81  | 81  | 81  |
| 82  | 82  | 82  | 82  |
| 83  | 83  | 83  | 83  |
| 84  | 84  | 84  | 84  |
| 85  | 85  | 85  | 85  |
| 86  | 86  | 86  | 86  |
| 87  | 87  | 87  | 87  |
| 88  | 88  | 88  | 88  |
| 89  | 89  | 89  | 89  |
| 90  | 90  | 90  | 90  |
| 91  | 91  | 91  | 91  |
| 92  | 92  | 92  | 92  |
| 93  | 93  | 93  | 93  |
| 94  | 94  | 94  | 94  |
| 95  | 95  | 95  | 95  |
| 96  | 96  | 96  | 96  |
| 97  | 97  | 97  | 97  |
| 98  | 98  | 98  | 98  |
| 99  | 99  | 99  | 99  |
| 100 | 100 | 100 | 100 |

PT

F.T. neu  
yy

Db 1148 GGTCTCAGCCTTGTGATGTTGGAGCATAGGTTGCAATTCCTATTGAATATTACCTTGTT 1207  
Qy 1045 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGCAATGATGGAACGAATATTAG 1104  
Db 1208 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGCAATGATGGAACGAATATTAG 1267  
Qy 1105 GACCCATACCAACACATGATTCGAAACACAGAAACGCAAGTATTTTCCACCATACC 1164  
Db 1268 GACCCATACCAACACATGATTCGAAACACAGAAACGCAAGTATTTTCCACCATACC 1327  
Qy 1165 AGCTAGATTGGGATGCAACACAGTCTCTGCTGTGATATGTTAGGAGACGCTGCAACCGT 1224  
Db 1328 AGCTAGATTGGGATGCAACACAGTCTCTGCTGTGATATGTTAGGAGACGCTGCAACCGT 1387  
Qy 1225 TGAAGGAATTTATGCTTTGTGATGATGAAGAACATGAGAAACTGTGTCACCTGGTTCGAA 1284  
Db 1388 TGAAGGAATTTATGCTTTGTGATGATGAAGAACATGAGAAACTGTGTCACCTGGTTCGAA 1447  
Qy 1295 GAATGTTAGATATGATCCAACTCAAGAAATACCTTGGATGAAGCATTCGAGCATCCTT 1344  
Db 1448 GAATGTTAGATATGATCCAACTCAAGAAATACCTTGGATGAAGCATTCGAGCATCCTT 1507  
Qy 1345 TCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTCTCT 1404  
Db 1508 TCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTCTCT 1567  
Qy 1405 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATCTTAATATTTTGTAAACATT 1464  
Db 1568 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATCTTAATATTTTGTAAACATT 1627  
Qy 1465 AAATATTTGTACAGTTAAGTGAATAATTTGTATGTTTGTATCAATPAGCATATTAAC 1524  
Db 1628 AAATATTTGTACAGTTAAGTGAATAATTTGTATGTTTGTATCAATPAGCATATTAAC 1687  
Qy 1525 TTGTTAAGCAAGTATGCTTGAATATCATTAGAAATTAATAATTTTCTTTT 1584  
Db 1688 TTGTTAAGCAAGTATGCTTGAATATCATTAGAAATTAATAATTTTCTTTT 1747  
Qy 1585 GAAATATACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGTGATTG 1644  
Db 1748 GAAATATACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGTGATTG 1807  
Qy 1645 ATCTTGCTTTTGTACATGAGGTCACCTCTGAGTGAATTTTGTGATGAAGAAAT 1704  
Db 1808 ATCTTGCTTTTGTACATGAGGTCACCTCTGAGTGAATTTTGTGATGAAGAAAT 1867  
Qy 1705 CTGACTACTTTATTTCTTAAAGAAATATCTTTATATATCTTCAAAATTTAGACTTAAC 1764  
Db 1868 CTGACTACTTTATTTCTTAAAGAAATATCTTTATATATCTTCAAAATTTAGACTTAAC 1927  
Qy 1765 TTTAAAGTTTCTTCTGTAAATTTTGAACGGGTGATATTAATTAACCTAGATAAGCA 1824  
Db 1928 TTTAAAGTTTCTTCTGTAAATTTTGAACGGGTGATATTAATTAACCTAGATAAGCA 1987  
Qy 1825 GGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTTATTAATTTTAAGTG 1884  
Db 1988 GGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTTATTAATTTTAAGTG 2047  
Qy 1895 TTGATTTCTTTTCACTGGGTGATGTCAGGTGATAACGACATTCATGGAAGGCATG 1944  
Db 2048 TTGATTTCTTTTCACTGGGTGATGTCAGGTGATAACGACATTCATGGAAGGCATG 2107  
Qy 1945 CAGTTTGTCCATGTCAGTGTGTTTAAATAAACACATACACACTTTTATTAAGATTA 2004  
Db 2108 CAGTTTGTCCATGTCAGTGTGTTTAAATAAACACATACACACTTTTATTAAGATTA 2167  
Qy 2005 AAATCTAACTGGAAGTCTGAGTGGAAATGACATTTTCCAAGTATGTTGGTGAAGTCA 2064  
Db 2168 AAATCTAACTGGAAGTCTGAGTGGAAATGACATTTTCCAAGTATGTTGGTGAAGTCA 2227  
Qy 2065 AGATATAAAATAGAAATTCGATGAGAGGTTTCAAGTTTAAATACCAAGTCTTAGAG 2124  
Db 2228 AGATATAAAATAGAAATTCGATGAGAGGTTTCAAGTTTAAATACCAAGTCTTAGAG 2287

2125 TCTTAAATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAAT 2184  
2288 TCTTAAATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAAT 2347  
2185 AAGTTTATTAATTAAGGCAATTTATGTCGTGATATTTCTTACGGGAGAAAGAGGATTTGA 2244  
2348 AAGTTTATTAATTAAGGCAATTTATGTCGTGATATTTCTTACGGGAGAAAGAGGATTTGA 2407  
2245 TTGAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAATTTGATTGTTA 2304  
2408 TTGAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAATTTGATTGTTA 2467  
2305 CATAAACTTTTGTACTTCA 2323  
2468 CATAAACTTTTGTACTTCA 2486

RESULT 5  
ADI29373  
ID ADI29373 standard; cDNA; 2488 BP.  
XX  
AC ADI29373;  
XX  
DT 22-APR-2004 (first entry)  
XX  
Human MARK3-associated cDNA #40.  
XX  
Human; ss; antisense gene therapy; MARK3;  
KW MAP/microtubule affinity-regulating kinase 3; cancer;  
KW Alzheimer's disease; neurodegenerative disorder;  
KW hyperproliferative disorder; cytostatic.  
XX  
Homo sapiens.  
XX  
US2003232771-A1.  
XX  
18-DEC-2003.  
XX  
17-JUN-2002; 2002US-00174319.  
XX  
17-JUN-2002; 2002US-00174319.  
XX  
(ISIS-) ISIS PHARM INC.  
PI Ward DT, Freier SM, Dobie KW;  
XX  
WPI; 2004-052188/05.  
DR P-FSDB; ADI29255.  
XX  
New antisense compound targeted to a nucleic acid molecule encoding microtubule-affinity-regulating kinases (MARK3), useful for modulating expression of MARK3 or for treating cancer or Alzheimer's disease.  
XX  
Disclosure; Fig 2; 233pp; English.  
XX  
The invention relates to a compound comprising a sequence comprising 8-80 base pairs (bp) targeted to a nucleic acid encoding MARK3 (MAP/microtubule affinity-regulating kinase 3), that specifically hybridizes with the nucleic acid encoding MARK3 and inhibits expression of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a composition comprising the compound and a carrier or diluent, inhibiting the expression of MARK3 in cells or tissues, treating an animal having or suspected of having a disease or condition associated with MARK3 and screening for an antisense compound. The antisense oligonucleotide is useful for preparing a composition for treating hyperproliferative disorder, particularly cancer and neurodegenerative diseases e.g. Alzheimer's disease. The present sequence is a MARK3 associated cDNA included in the figures but not mentioned anywhere else in the specification.  
XX  
Sequence 2488 BP; 819 A; 400 C; 504 G; 764 T; 0 U; 1 Other;  
SQ

|   |      |   |      |    |      |  |      |  |  |
|---|------|---|------|----|------|--|------|--|--|
| Query Match 87.0%; Score 2048; DB 12; Length 2488;<br>Best Local Similarity 95.9%; Pred. No. 0;<br>Matches 2148; Conservative 0; Mismatches 1; Indels 90; Gaps 1; |      |   |      |    |      |  |      |  |  |
| QY  | 85   | GTCAATATTAGAACCAAGGTCCTTGAATGAGCGAGATATTCGGACCGGAGATACGTTG      | 144  | DB | 1165 | AGCTAGATTGGGATGAACACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT  | 1224 |  |  |
| DB  | 338  | GTCAATATTAGAACCAAGGTCCTTGAATGAGCGAGATATTCGGACCGGAGATACGTTG      | 397  | QY | 1328 | AGCTAGATTGGGATGAACACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT  | 1387 |  |  |
| QY  | 145  | ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGAGACATTG   | 204  | QY | 1225 | TGAAGGAATTTATGCTTTTGTATGATGAAGAACATCAGAAACTGTTTGACCTGGTTCGAA     | 1284 |  |  |
| DB  | 398  | ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGAGACATTG   | 457  | DB | 1388 | TGAAGGAATTTATGCTTTTGTATGATGAAGAACATCAGAAACTGTTTGACCTGGTTCGAA     | 1447 |  |  |
| QY  | 205  | AAAGCGGTATCGAATCCATGCGAGTAATCTTCAGTCCGACGAGGAGACGATCCTTA        | 264  | QY | 1285 | GAATGTTAGAAATATGATCCAACTCAAAGAAATTAACCTTGATGAAGCAATTCGACGATCCTT  | 1344 |  |  |
| DB  | 458  | AAAGCGGTATCGAATCCATGCGAGTAATCTTCAGTCCGACGAGGAGACGATCCTTA        | 517  | DB | 1448 | GAATGTTAGAAATATGATCCAACTCAAAGAAATTAACCTTGATGAAGCAATTCGACGATCCTT  | 1507 |  |  |
| QY  | 265  | AAAGGAAGCGCAATAGACACACTGTTCAAGTCATCAGTCAGTTCGAAAGACCCAGGAAGA    | 324  | QY | 1345 | TCCTTGACCTTATTAATAAGAAATGAAATGGAATCAGTGGTCTTACTATATATCTTCTCT     | 1404 |  |  |
| DB  | 518  | AAAGGAAGCGCAATAGACACACTGTTCAAGTCATCAGTCAGTTCGAAAGACCCAGGAAGA    | 561  | DB | 1508 | TCCTTGACCTTATTAATAAGAAATGAAATGGAATCAGTGGTCTTACTATATATCTTCTCT     | 1567 |  |  |
| QY  | 325  | AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACG     | 384  | QY | 1405 | AGAAGAGATTACTTAAAGACTGTGTCAGTCAACTAAACATTTCTAAATATTTTGTGAAACATT  | 1464 |  |  |
| DB  | 562  | -----   | 561  | DB | 1568 | AGAAGAGATTACTTAAAGACTGTGTCAGTCAACTAAACATTTCTAAATATTTTGTGAAACATT  | 1627 |  |  |
| QY  | 385  | TTCTAAGACGAAGATATAAATCGTGACACTTTGGGTGAAGAGCCCTTTGGCAAGTTG       | 444  | QY | 1465 | AAATTAATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC        | 1524 |  |  |
| DB  | 562  | -----NATGAATCGTGACACTTTGGGTGAAGAGCCCTTTGGCAAGTTG                | 607  | DB | 1628 | AAATTAATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC        | 1687 |  |  |
| QY  | 445  | TAGAGTGCATTGATCATGGCATGGCATGCGATGCGATGTAGCAGTGAANAATCGTAAANAATG | 504  | QY | 1525 | TTGTTAAGCAAGTATCGTCTTGATATGCAATTAGAAAAATTAATAATTTTCTTTT          | 1584 |  |  |
| DB  | 608  | TAGAGTGCATTGATCATGGCATGGCATGCGATGCGATGTAGCAGTGAANAATCGTAAANAATG | 667  | DB | 1688 | TTGTTAAGCAAGTATCGTCTTGATATGCAATTAGAAAAATTAATAATTTTCTTTT          | 1747 |  |  |
| QY  | 505  | TAGGCCGTATACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA | 564  | QY | 1585 | GAAATTTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATTG     | 1644 |  |  |
| DB  | 668  | TAGGCCGTATACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA | 727  | DB | 1748 | GAAATTTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATTG     | 1807 |  |  |
| QY  | 565  | CTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAAATGTTGATCATCATGTC        | 624  | QY | 1645 | ATCTTGCCCTTTTGTACATGAGAGTCACTCTGAAAGTGAATTTTGTGAGTGAAGGAAAT      | 1704 |  |  |
| DB  | 728  | CTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAAATGTTGATCATCATGTC        | 787  | DB | 1808 | ATCTTGCCCTTTTGTACATGAGAGTCACTCTGAAAGTGAATTTTGTGAGTGAAGGAAAT      | 1867 |  |  |
| QY  | 625  | ATGTTTGTATTGTTTGAATCTAGGACTTAGTACTTACGATTTCAATTAAGAAACA         | 684  | QY | 1705 | CTTGACCTACTTTATATCTTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAAC   | 1764 |  |  |
| DB  | 788  | ATGTTTGTATTGTTTGAATCTAGGACTTAGTACTTACGATTTCAATTAAGAAACA         | 847  | DB | 1868 | CTTGACCTACTTTATATCTTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAAC   | 1927 |  |  |
| QY  | 685  | GCTTCTGCCATTTCAAATTTGACCAATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA     | 744  | QY | 1765 | TTTAAAGTTTTTCTCTGTAAATTTGTTGAACGGGTGATTATTTAACTCTAGATAAGCA       | 1824 |  |  |
| DB  | 848  | GCTTCTGCCATTTCAAATTTGACCAATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA     | 907  | DB | 1928 | TTTAAAGTTTTTCTCTGTAAATTTGTTGAACGGGTGATTATTTAACTCTAGATAAGCA       | 1987 |  |  |
| QY  | 745  | TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCTGAAAATATTTTGT     | 804  | QY | 1825 | GGTACTAGAAAACCAAACTCAGAAAAATGTTTACTGTTGAGAAATCTATTTAAATTTAAGTG   | 1884 |  |  |
| DB  | 908  | TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCTGAAAATATTTTGT     | 967  | DB | 1988 | GGTACTAGAAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTATTTAAATTTAAGTG    | 2047 |  |  |
| QY  | 805  | TTGTGAAGTCTGACTATGATCAATATAAATTTCTAAATGAAGCTGATGAACGCACAC       | 864  | QY | 1885 | TTGTATTTCTTTTTCATTTGGGTGATGTCAGGGTGATAACCCAGACATTCATGGAAGGCGATG  | 1944 |  |  |
| DB  | 968  | TTGTGAAGTCTGACTATGATCAATATAAATTTCTAAATGAAGCTGATGAACGCACAC       | 1027 | DB | 2048 | TTGTATTTCTTTTTCATTTGGGTGATGTCAGGGTGATAACCCAGACATTCATGGAAGGCGATG  | 2107 |  |  |
| QY  | 865  | TGAATAACACAGATATCAAAGTTGTTGACTTTTGAAGTGCAACGATATGATGAACATC      | 924  | QY | 1945 | CAGTTTGTCCATTGTGACAGTTTGTTTAATAAAAAACACATACACACATTTTAAAGATTA     | 2004 |  |  |
| DB  | 1028 | TGAATAACACAGATATCAAAGTTGTTGACTTTTGAAGTGCAACGATATGATGAACATC      | 1087 | DB | 2108 | CAGTTTGTCCATTGTGACAGTTTGTTTAATAAAAAACACATACACACATTTTAAAGATTA     | 2167 |  |  |
| QY  | 925  | ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT   | 984  | QY | 2005 | AAATCTAACTGGAAAGTCAGCTTGGAAAAATGGACATTTCCAAAGTATGTTTGGTGGTGCAC   | 2064 |  |  |
| DB  | 1088 | ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT   | 1147 | DB | 2168 | AAATCTAACTGGAAAGTCAGCTTGGAAAAATGGACATTTCCAAAGTATGTTTGGTGGTGCAC   | 2227 |  |  |
| QY  | 985  | GGTCTCAGCCTTGTGATGTTTGGAGCATAGTTTGCATCTTATTTGAATATTTACCTTGGTT   | 1044 | QY | 2065 | AGATATAAAATAGAAATTTCTGATGAGAGTTTTCAGTTTTTAAATACCAAGTCTTTAGGAG    | 2124 |  |  |
| DB  | 1148 | GGTCTCAGCCTTGTGATGTTTGGAGCATAGTTTGCATTTCTTATTTGAATATTTACCTTGGTT | 1207 | DB | 2228 | AGATATAAAATAGAAATTTCTGATGAGAGTTTTCAGTTTTTAAATACCAAGTCTTTAGGAG    | 2287 |  |  |
| QY  | 1045 | TCACAGTCTTTTCAGACTCATGATAGTAAGAGCACCCTGGCAATGATGGAACCAATATTAG   | 1104 | QY | 2125 | TCCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTTAAAGAAAT | 2184 |  |  |
| DB  | 1208 | TCACAGTCTTTTCAGACTCATGATAGTAAGAGCACCCTGGCAATGATGGAACCAATATTAG   | 1267 | DB | 2288 | TCCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTTAAAGAAAT | 2347 |  |  |
| QY  | 1105 | GACCCATACCAACACATGATTCAGAAAAACGCAAGTATTTTCCACCATAAAC            | 1164 | QY | 2185 | AAGTTTATTAATTAGGCAATTTATGTCGTGATTAATCTTACGGGAGAAAGAGGATTTGA      | 2244 |  |  |

Db 2348 AAGTTTATTATAGGCAATTTATGCTGTGTAATTTCTTACGGAGAAAGGATTGCA 2407  
 Qy 2245 TTGGAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCGAAATTTAATTGTTGTTA 2304  
 Db 2408 TTGGAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCGAAATTTAATTGTTGTTA 2467  
 Qy 2305 CATAAACATTTTGACTTCA 2323  
 Db 2468 CATAAACATTTTGACTTCA 2486  
 RESULT 6  
 ADP24393  
 ID ADP24393 standard; cDNA; 1814 BP.  
 AC ADP24393;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE PRO polypeptide encoding cDNA SEQ ID NO:1571.  
 XX  
 KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
 KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
 KW antipruritic; antiallergic; antiasthmatic; hepatotropic; respiratory;  
 KW gene therapy; immune system.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2004041170-A2.  
 XX  
 XX 21-MAY-2004.  
 PD  
 XX 30-OCT-2003; 2003WO-US034312.  
 PF  
 XX 01-NOV-2002; 2002US-0423394P.  
 PR  
 XX (GETH ) GENENTECH INC.  
 FA  
 XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
 PI Wu TD;  
 PI  
 XX WPI; 2004-419628/39.  
 DR P-PSDB; ADP24394.  
 DR  
 XX  
 PT New PRO polypeptides and polynucleotides, useful for treating e.g.  
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
 PT renal disease, or demyelinating diseases of the central or peripheral  
 PT nervous system.  
 PT  
 PS Claim 1; SEQ ID NO 1571; 2940pp; English.  
 XX  
 CC The invention relates to a novel isolated nucleic acid and the PRO  
 CC polypeptide encoded by it. A protein of the invention has  
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
 CC osteopathic, antidiabetic, dermatological, antipruritic, antiallergic,  
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
 CC agonist, antagonist, or antibody that specifically binds to the  
 CC polypeptide is useful for treating an immune related disorder such as  
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
 CC disease, a demyelinating disease of the central or peripheral nervous  
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
 CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
 CC disease, infectious or autoimmune chronic active hepatitis, primary  
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food

CC hypersensitivity, urticaria, an immunologic disease of the lung,  
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
 CC pneumonitis, a transplantation associated disease, graft rejection or  
 CC graft-versus-host disease. The present sequence encodes a PRO protein of  
 CC the invention.  
 XX  
 SQ Sequence 1814 BP; 609 A; 299 C; 385 G; 520 T; 0 U; 1 Other;  
 Query Match 69.3%; Score 1632; DB 13; Length 1814;  
 Best Local Similarity 99.3%; Pred. No. 5.8e-312;  
 Matches 1638; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 Qy 85 GTCATTATTAGAACCAAGTCTTGAATGACGAGATATTCGGACCGGAGATACGTTG 144  
 Db 161 GTCATTATTAGAACCAAGTCTTGAATGACGAGATATTCGGACCGGAGATACGTTG 220  
 Qy 145 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGAGACATTG 204  
 Db 221 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGAGACATTG 280  
 Qy 205 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAGAGACAGTCTCTA 264  
 Db 281 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAGAGACAGTCTCTA 340  
 Qy 265 AAAGGAGCGCAATAGACACATGTTCAAGTCACTAGTTCAGTTCGAGAGACCCGAGGA 324  
 Db 341 AAAGGAGCGCAATAGACACATGTTCAAGTCACTAGTTCAGTTCGAGAGACCCGAGGA 400  
 Qy 325 AAAGTCCAGGATATGAGGATGATGAGGAGGTCCACTGATCTGTCAGAGTGAGACG 384  
 Db 401 AAAGTCCAGGATATGAGGATGATGAGGAGGTCCACTGATCTGTCAGAGTGAGACG 460  
 Qy 385 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 444  
 Db 461 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 520  
 Qy 445 TAGAGTGCATTGATCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGG 504  
 Db 521 TAGAGTGCATTGATCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGG 580  
 Qy 505 TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 564  
 Db 581 TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 640  
 Qy 565 CTGATCCCAATAGTGTCTTCGAGATGTGTCCAGATGTGTAGAAATGGTTGATCATCATGGTC 624  
 Db 641 CTGATCCCAATAGTGTCTTCGAGATGTGTCCAGATGTGTAGAAATGGTTGATCATCATGGTC 700  
 Qy 625 ATGTTTGTATTGTTGTTGAACTACTGGGACTTTAGTACTTTACGATTTTCAATTAAGAAAACA 684  
 Db 701 ATGTTTGTATTGTTGTTGAACTACTGGGACTTTAGTACTTTACGATTTTCAATTAAGAAAACA 760  
 Qy 685 GCTTTCTGCCATTTCAAAATTCACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 744  
 Db 761 GCTTTCTGCCATTTCAAAATTCACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 820  
 Qy 745 TAAATTTTATCATCAATAAATTAACCCATACAGATCTCGAAGCCTGAAATATTTTGT 804  
 Db 821 TAAATTTTATCATCAATAAATTAACCCATACAGATCTCGAAGCCTGAAATATTTTGT 880  
 Qy 805 TTGTGAAGTCTGATGTAGTCAATAATAATTTCAAAATGAAGAGTGAAGAGTGAAGAGTGA 864  
 Db 881 TTGTGAAGTCTGATGTAGTCAATAATAATTTCAAAATGAAGAGTGAAGAGTGAAGAGTGA 940  
 Qy 865 TGAAAAACACAGATATCAAGTCTGTGACTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 924  
 Db 941 TGAAAAACACAGATATCAAGTCTGTGACTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1000  
 Qy 925 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCTCCGAGGTCAATTTTGGCTTTAGGTT 984  
 Db 1001 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCTCCGAGGTCAATTTTGGCTTTAGGTT 1060  
 Qy 995 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTGTCATTTCTTATTGAATATTACCTTGGTT 1044









Db 1. TTCAATTTGACCATCATCAGGAGATGGGTATCAGATCGCCAGTCAATAAATTTTTTAC 60  
Qy 757 ATCAATAAATAAATTAACCATACAGATCAGATCGGAGCCGTGAATAATATTTTGTGTGAAGTCTG 816  
Db 61 ATCAATAAATAAATTAACCATACAGATCAGATCGGAGCCGTGAATAATATTTTGTGTGAAGTCTG 120  
Qy 817 ACTATGTAGTCAGATTAATAATTTCTAAATGAACGTGATGAACGACACACTGAAAAACACAG 876  
Db 121 ACTATGTAGTCAGATTAATAATTTCTAAATGAACGTGATGAACGACACACTGAAAAACACAG 180  
Qy 877 ATATCAAAAGTTGTGACATTTTGGAAATGCAACGTATGATGAACATCACAGTACTTTGG 936  
Db 181 ATATCAAAAGTTGTGACATTTTGGAAATGCAACGTATGATGAACATCACAGTACTTTGG 240  
Qy 937 TGTCTACCCGACATACAGAGCTCCGAGGTCATTTTGGCTTTAGTGTGCTCAGCCTT 996  
Db 241 TGTCTACCCGACATACAGAGCTCCGAGGTCATTTTGGCTTTAGTGTGCTCAGCCTT 300  
Qy 997 GTGATGTTTGGAGCATAGTTTGCATTTCTATTGAATATTAATCTTGGTTTTCAGTCTTTTC 1056  
Db 301 GTGATGTTTGGAGCATAGTTTGCATTTCTATTGAATATTAATCTTGGTTTTCAGTCTTTTC 360  
Qy 1057 AGACTCATGATAGTAAAGAGACACTGGCAATGATGAACGAATATTAGGACCCATACAC 1116  
Db 361 AGACTCATGATAGTAAAGAGACACTGGCAATGATGAACGAATATTAGGACCCATACAC 420  
Qy 1117 AACACATGATTCAGAAAAACAGAAACCGAATATTTTCCATTAACCAACAGCTAGATTGG 1176  
Db 421 AACACATGATTCAGAAAAACAGAAACCGAATATTTTCCATTAACCAACAGCTAGATTGG 480  
Qy 1177 ATGAACACAGTTCTGCTGTAGATATGTTAGGACGCTGCAAAACCGTTGAAGAAATTTA 1236  
Db 481 ATGAACACAGTTCTGCTGTAGATATGTTAGGACGCTGCAAAACCGTTGAAGAAATTTA 540  
Qy 1237 TGCCTTTGTATGATGAAGAACATGAGAACTGTTTGAACCTGTTTGAAGAAATTTAGAT 1296  
Db 541 TGCCTTTGTATGATGAAGAACATGAGAACTGTTTGAACCTGTTTGAAGAAATTTAGAT 600  
Qy 1297 ATGATCCAACTCAAGAAATTAACCTTGGATGAGCAATTCAGCATCTTTCTTTGACTTAT 1356  
Db 601 ATGATCCAACTCAAGAAATTAACCTTGGATGAGCAATTCAGCATCTTTCTTTGACTTAT 660  
Qy 1357 TAAAAAGAAATGAATGGAAATCAGTGTCTTACTATATCTCTCTAGAGAGATTAC 1416  
Db 661 TAAAAAGAAATGAATGGAAATCAGTGTCTTACTATATCTCTCTAGAGAGATTAC 720  
Qy 1417 TTAAGACTGTGTGAGTCAAACTAAACATTTCTAATATTTTGTAAACATTAATTTTGT 1476  
Db 721 TTAAGACTGTGTGAGTCAAACTAAACATTTCTAATATTTTGTAAACATTAATTTTGT 780  
Qy 1477 ACAGTTTAACTGTAATATTTGATGTTTGTATCAATAGCATTAATTAACCTGTTAAGCAAG 1536  
Db 781 ACAGTTTAACTGTAATATTTGATGTTTGTATCAATAGCATTAATTAACCTGTTAAGCAAG 840  
Qy 1537 TATGCTCTTGTATAATGATAGAAAAATTAATAATTTTCTTTTGAATTAACCAT 1596  
Db 841 TATGCTCTTGTATAATGATAGAAAAATTAATAATTTTCTTTTGAATTAACCAT 900  
Qy 1597 TTTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATATGATGATGATGATGATGATGATGAT 1656  
Db 901 TTTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATATGATGATGATGATGATGATGATGAT 960  
Qy 1657 GTACATGAGGTCACCTCTGAGTGTATTTTGTGATGATGATGATGATGATGATGATGATGAT 1716  
Db 961 GTACATGAGGTCACCTCTGAGTGTATTTTGTGATGATGATGATGATGATGATGATGATGAT 1020  
Qy 1717 ATATTTCTTAAAGGAATATTTCTTTTATATATCTTCAAAATTTGAACCTTTAAAGTTTT 1776  
Db 1021 ATATTTCTTAAAGGAATATTTCTTTTATATATCTTCAAAATTTGAACCTTTAAAGTTTT 1080  
Qy 1777 TCTTCTGTAATTTGTTGAACGGGTGATATTTAATCTTAGATAAGCAGTACTAGAAAC 1836  
Db 1081 TCTTCTGTAATTTGTTGAACGGGTGATATTTAATCTTAGATAAGCAGTACTAGAAAC 1140

Qy 1837 CAAAACCTCAGAAAATGTTTACTGTAGAAATTTTAAATTTTAAAGTGTGTATCTTTT 1896  
Db 1141 CAAAACCTCAGAAAATGTTTACTGTAGAAATTTTAAATTTTAAAGTGTGTATCTTTT 1200  
Qy 1897 TCATTGGGTGATGTCAGGGTGATAACAGACATTCATGGAAGGCATGCGAGTTTGTCCAT 1956  
Db 1201 TCATTGGGTGATGTCAGGGTGATAACAGACATTCATGGAAGGCATGCGAGTTTGTCCAT 1260  
Qy 1957 TGTGACAGATTTGTTTAAATAAACAACATACACACTTTTAAATTAAGATTAAATCTAACTGG 2016  
Db 1261 TGTGACAGATTTGTTTAAATAAACAACATACACACTTTTAAATAAATAAATAAATAAATAA 1320  
Qy 2017 AAA 2019  
Db 1321 AAA 1323  
RESULT 9  
AD007809  
ID AD007809 standard; cDNA; 1446 BP.  
XX AD007809;  
AC AD007809;  
XX 01-JUL-2004 (first entry)  
XX Human polynucleotide #39.  
XX Human; gene; ss; fat cell number; fat cell size; obesity; diabetes;  
XX anorectic; antidiabetic.  
XX Homo sapiens.  
XX US2004071700-A1.  
XX 15-APR-2004.  
XX 09-OCT-2002; 2002US-00267502.  
XX 09-OCT-2002; 2002US-00267502.  
XX (LIFE-) LIFE SCI DEV CORP.  
XX Kim J, Galant R;  
XX WPI; 2004-328526/30.  
XX P-PSDB; ADO08026.  
XX Identifying compounds that influence fat cell number or size for treating  
or preventing obesity or diabetes by exposing the cell to the agent and  
identifying fat cell number or size relative to cells not exposed to the  
agent.  
XX Claim 1; SEQ ID NO 135; 275pp; English.  
XX The invention relates to a method of identifying compounds that influence  
fat cell number or size comprising providing a cell that expresses a gene  
and an agent, exposing the cell to the agent and identifying fat cell  
number or size relative to cells not exposed to the agent. The method  
also comprises providing an expression vector and an agent, exposing the  
vector to the agent, detecting a change in expression of the gene  
relative to expression of the gene in an expression vector not exposed to  
the agent, treating a subject with the agent and identifying fat cell  
number or size in the subject. The agent comprises a mammal, preferably a human. The  
oligonucleotide. The subject comprises a polypeptide and an agent, exposing the  
method also comprises providing a polypeptide of the agent to the  
polypeptide or a change in an activity of the polypeptide, treating a  
subject with the agent and identifying fat cell number or size in the  
subject. The agent comprises an antibody. A method of regulating fat cell  
number or size comprises providing a subject containing fat cells and an  
agent that changes the expression of a gene, and treating the subject  
with the agent under conditions so that fat cell size or number in the

CC subject is altered. The method is useful for identifying compounds that  
 CC influence fat cell number or size, for preparing a composition for  
 CC treating or preventing obesity or diabetes. This sequence represents  
 CC human cDNA used in the scope of the invention.  
 XX  
 SQ Sequence 1446 BP; 474 A; 258 C; 332 G; 382 T; 0 U; 0 Other;

Query Match 54.6%; Score 1286; DB 12; Length 1446;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-244;  
 Matches 1286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |      |  |      |
|----|------|--|------|
| QY | 85   | GTCAATATTAGAACCAAGGCTCTTGAATGACGAGATATTCGGACCGGAGATACGTTG    | 144  |
| DB | 161  | GTCAATATTAGAACCAAGGCTCTTGAATGACGAGATATTCGGACCGGAGATACGTTG    | 220  |
| QY | 145  | ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTCACAGACATTTG    | 204  |
| DB | 221  | ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTCACAGACATTTG    | 280  |
| QY | 205  | AAAGCGGTATCGAATCCATCGCAGTAATCTTCAGTCCGACGAGGAGACGTCCTTA      | 264  |
| DB | 281  | AAAGCGGTATCGAATCCATCGCAGTAATCTTCAGTCCGACGAGGAGACGTCCTTA      | 340  |
| QY | 265  | AAAGGAAGCGCAATAGACACTGTTCAAGTCATCATGTCAGTTCGAAGACCCACGAAGA   | 324  |
| DB | 341  | AAAGGAAGCGCAATAGACACTGTTCAAGTCATCATGTCAGTTCGAAGACCCACGAAGA   | 400  |
| QY | 325  | AAAGATCCAGGAGTATAGAGGATGATGAGGAGTCACTGATCTGTCAAAGTGAGACG     | 384  |
| DB | 401  | AAAGATCCAGGAGTATAGAGGATGATGAGGAGTCACTGATCTGTCAAAGTGAGACG     | 460  |
| QY | 385  | TTCTAAGAGCAAGATPATGAATCGTGGACACTTTTGGGTGAAGAGCCCTTTGGCAAGTTG | 444  |
| DB | 461  | TTCTAAGAGCAAGATPATGAATCGTGGACACTTTTGGGTGAAGAGCCCTTTGGCAAGTTG | 520  |
| QY | 445  | TAGATGATTCATGATGGCATGGATGGATGATGATGATGATGATGATGATGATGATGATG  | 504  |
| DB | 521  | TAGATGATTCATGATGGCATGGATGGATGATGATGATGATGATGATGATGATGATGATG  | 580  |
| QY | 505  | TAGGCGGTACCGTCAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAATAGTA  | 564  |
| DB | 581  | TAGGCGGTACCGTCAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAATAGTA  | 640  |
| QY | 565  | CTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGTTGATCATCATGTC    | 624  |
| DB | 641  | CTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGTTGATCATCATGTC    | 700  |
| QY | 625  | ATGTTTGTATTGTTTGAATCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA  | 684  |
| DB | 701  | ATGTTTGTATTGTTTGAATCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA  | 760  |
| QY | 685  | GCTTTCGCCATTTCAAAATTTGACACATCAGGACATGCGGTATCAGATCTGCCAGTCAA  | 744  |
| DB | 761  | GCTTTCGCCATTTCAAAATTTGACACATCAGGACATGCGGTATCAGATCTGCCAGTCAA  | 820  |
| QY | 745  | TAAATTTTTTACATCAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT    | 804  |
| DB | 821  | TAAATTTTTTACATCAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT    | 880  |
| QY | 805  | TTGTGAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG | 864  |
| DB | 881  | TTGTGAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG | 940  |
| QY | 865  | TGAAAAACACAGATATCAAGTGTGTTGACTTTTGAAGTGCACGTATGATGATGATGATG  | 924  |
| DB | 941  | TGAAAAACACAGATATCAAGTGTGTTGACTTTTGAAGTGCACGTATGATGATGATGATG  | 1000 |
| QY | 925  | ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGTT   | 984  |
| DB | 1001 | ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGTT   | 1060 |
| QY | 985  | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGATTTCTTATTGAATATTACCTTGGTT | 1044 |

|    |      |  |      |
|----|------|--|------|
| DB | 1061 | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGATCTTTTGAATATTACCTTGGTT      | 1120 |
| QY | 1045 | TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGGAACGAATATTAG   | 1104 |
| DB | 1121 | TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGGAACGAATATTAG   | 1180 |
| QY | 1105 | GACCCATACCAACACACATGATTTCAAGAAACGAAGAAACGCAAGTATTTCACCAATACC   | 1164 |
| DB | 1181 | GACCCATACCAACACACATGATTTCAAGAAACGAAGAAACGCAAGTATTTCACCAATACC   | 1240 |
| QY | 1165 | AGCTAGATTGGGATGAACACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGCAACCGT    | 1224 |
| DB | 1241 | AGCTAGATTGGGATGAACACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGCAACCGT    | 1300 |
| QY | 1225 | TGAAGGAATTTATGCTTTTGTCTATGATGAAGAACATGAGAAACTGTTGACCTGGTTGAA   | 1284 |
| DB | 1301 | TGAAGGAATTTATGCTTTTGTCTATGATGAAGAACATGAGAAACTGTTGACCTGGTTGAA   | 1360 |
| QY | 1285 | GAATGTTAGAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGACGATCCCTT | 1344 |
| DB | 1361 | GAATGTTAGAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGACGATCCCTT | 1420 |
| QY | 1345 | TCCTTGACTTATTAAGAAAGAAATGA                                     | 1370 |
| DB | 1421 | TCCTTGACTTATTAAGAAAGAAATGA                                     | 1446 |

RESULT 10  
 ADI26140  
 ID ADI26140 standard; cDNA; 1881 BP.  
 XX  
 AC ADI26140;  
 XX  
 DT 22-APR-2004 (first entry)  
 XX  
 DE Human cDNA encoding protein that promotes STAT6 activation #53.  
 XX  
 KW ss; gene; human; signal transducer and activator of transcription 6;  
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;  
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;  
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;  
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;  
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003104277-A2.  
 XX  
 PD 18-DEC-2003.  
 XX  
 PF 05-JUN-2003; 2003WO-JP007123.  
 XX  
 PR 05-JUN-2002; 2002JP-00164257.  
 PR 06-JUN-2002; 2002US-0385912P.  
 PR 26-DEC-2002; 2002JP-00377326.  
 PR 27-DEC-2002; 2002US-0436467P.  
 PR 15-MAY-2003; 2003JP-00137505.  
 PR 16-MAY-2003; 2003US-0470836P.  
 XX  
 PA (ASAH ) ASAH KASEI KK.  
 XX  
 PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Iehizawa K;  
 XX  
 DR WPI; 2004-122214/12.  
 DR P-FSDB; ADI26141.  
 XX  
 PT New signal transducer and activator of transcription 6 activation  
 PT promoting purified protein, for diagnosing and treating disease  
 PT associated with activation/inhibition of transcription factor e.g.  
 PT diabetes and cancer.  
 XX  
 PS Claim 4; SEQ ID NO 105; 1368pp; English.  
 XX

The invention relates to a purified protein promoting signal transducer and activator of transcription 6 activation (STAT6). The protein is useful for the producing an antibody, which involves administering the protein or its epitope-bearing fragments to a non-human animal as an antigen. The nucleic acid is useful for diagnosing a disease or susceptibility to a disease related to expression or activity of the protein. A transformant expressing the protein is useful for screening compounds which inhibit or promote STAT6 activation. A transformant expressing the protein is useful for producing a pharmaceutical composition. Compositions, antibodies and antisense molecules are useful for the treating a disease associated with STAT6 activation such as allergic diseases, inflammation, autoimmune diseases, diabetes, hyperlipidaemia, infectious disease and cancers. Compositions are useful for treating disease associated with STAT6 activation and/or prevention of Th1 hyperactive diseases. Compositions are also useful in rheumatoid arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma, allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage, viral hepatitis and AIDS. The protein has efficient promoting STAT6 activity. The protein or nucleic acid is effectively useful for screening compounds for treating and preventing disease associated with excessive activation or inhibition of STAT6. The present sequence represents a human cDNA encoding a protein which promotes STAT6 activation.

Sequence 1881 BP; 570 A; 344 C; 450 G; 517 T; 0 U; 0 Other;  
Query Match 54.3%; Score 1277.4; DB 12; Length 1881;  
Best Local Similarity 87.9%; Pred No. 4.6e-242;  
Matches 1413; Conservative 0; Mismatches 193; Indels 3; Gaps 2;

85 GTCATTATTAGAACGAAGGCTCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 144  
250 GTCATTATTAGAACGAAGGCTCTTGAATGACGAGATTATCGGACCGGAGATACATTG 309  
145 ACGAATACAGAAATGACTCTGTGAGGATATGTTCTTAGACATTATCACAGACATTG 204  
310 ATGAATACAGAAATGACTCTGTGAGGATATGTTCTTAGACATTATCACAGACATTG 369  
205 AAAGCGGATATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAGAGACGTCCTA 264  
370 AAAGCACTTACCGGATCCATTGTCAGTAAATCTTCAGTCCGACGAGGAGAGACGTCCTA 429  
265 AAAGGAAGCGAAATAGACATCTGTTCAAGTCATCATGTCAGTCCGAGGAGACCGAAGGA 324  
430 AGAGAAGCGTAATAGACCCCTGTGCAAGTCATCATGTCAGTCCGAGGAGACCGAAGGA 489  
325 AAAGATCCAGAGTATACAGATGATGAGGAGGTCACCTGATCTGTCAAAAGTGAGAGCG 384  
490 AAAGATCCAGAGTATACAGAGTATGAGGAGGTCACCTGATCTGTCAAAAGTGAGAGCG 549  
385 TTCTAAGAGCAAGATATGAATATCGTGACATCTTGGGTGAAGGAGCCCTTGGCAAAAGTTG 444  
550 TTCTAAGAGCAAGATATGAATATCGTGACATCTTGGGTGAAGGAGCCCTTGGCAAAAGTTG 609  
445 TAGAGTGCATTGATCATCGGCATGGATGCGATGTCATGATGAGAGTGAATAATGATAAATG 504  
610 TAGAGTGCATTGATCATCGGCATGGATGCGTTCATGATGAGAGTGAATAATGATAAATG 669  
505 TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAGATATGAGGACCTTAATAGTA 564  
670 TAGGCGGTACCGGAGGAGCAGCTCGTTCGAAATCCAGATATGAGGACCTTGAACAGCA 729  
565 CTGATCCCAATAGTGTCTTCGATGTCCTGAGATGCTAGAAATGTTTTCATCATGATGTC 624  
730 CTGACCCCAACAGTGTCTTCGATGTCCTGAGATGCTAGAAATGTTTTCATCATGATGTC 789  
625 ATGTTTGTATTGTTGAACTATCGGACCTTAGTATCTTACATGATTTTCAATTAAGAAACA 684  
790 ATGTTTGTATTGTTGTTGAGCTGTGGGACCTTAGTATCTTACATGATTTTCAATTAAGAAACA 849  
685 GCTTCTGCCATTTCAAAATGACCAATCAGGAGATGCGGATCATGATCTGCCAGTCAA 744  
850 GTTTTCTGCCATTTCAAAATGATCAATCAGGCAAAATGGCTTATCAGATCTGCCAGTCTA 909

Qy 745 TAAATTTTTCATCATATAATAATAAATTAACCCCATACAGATCTGAAGCCCTGAAATAATTTTGT 804  
Db 910 TAAATTTTTCATCATATAATAATAAATTAACACACACGACCTAAAAACCTGAAATAATTTTAT 969  
Qy 805 TTGTGAAGTCTGACTATGTAGTCAAAATATATTTCTAAATGAAACGCTGATGAACGACAC 864  
Db 970 TTGTGAAGTCTGACTATGTAGTCAAAATATATTTCTAAATGAAACGAGATGAGGCGACAT 1029  
Qy 865 TGAATAACACAGATATCAAAAGTGTGTGACTTTTGAAGTGCAACGATGATGATGAACATC 924  
Db 1030 TGAATAACACAGATATCAAAAGTGTGTGACTTTTGAAGTGCAACATATGACGACGACATC 1089  
Qy 925 ACAGTACTTTGGTGTCTACCCGGCCTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 984  
Db 1090 ATAGTACTTTGGTGTCTCCACAGGCTACAGGCTCCAGAGGTCAATTTTGGCTCTAGGTT 1149  
Qy 985 GGTCTAGCCTTTGTGATGTTTGGAGCATAGTGTGATCTTATTTGAATATTTACCTTGGTT 1044  
Db 1150 GGTCTAGCCTTTGTGATGTTTGGAGCATAGGCTGCTATTTTATTTGAGTACTACCTTGGT 1209  
Qy 1045 TCAGAGTCTTTGAGACTCATGATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAG 1104  
Db 1210 TCAGAGTCTTTGAGACCCACGATAGTAAAGAGCACCTGGCAATGATGGAAGCGATCTAG 1269  
Qy 1105 GACCCATACCAACACATGATTTCCAGAAACCAAGAAACGCAAGTATTTTCCACATAACC 1164  
Db 1270 GACCCATCCAGCAGCATATGATCCAGAAACCAAGAAACGCAAGTATTTTCCACATAACC 1329  
Qy 1165 AGCTAGATTGGGATGAACACAGTTCTGCTGTAGATATGTTAGGAGCGCTGCAACCGT 1224  
Db 1330 AGCTAGATTGGGAGCAGCATAGTTTCACTGGGAGATATGTTAGGAGCGCTGCAACCGT 1389  
Qy 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTGCACCTGGTTGAA 1284  
Db 1390 TAAAGGAATTTATGCTGTGTCATGACGAGAGCATGAGAACTGTTTGCACCTGGTTGAA 1449  
Qy 1285 GAATTTTAGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATGTGAGCATCCTT 1344  
Db 1450 GAATTTTGGAGTATGATCCAGCAGGAGAGGATCACTTGGATGAAGCATGTGAGCATCCTT 1509  
Qy 1345 TCTTTGACTTATTAAGAAAGAAATGAATGGGAATCACTGCTTACTATATATCTTCTCT 1404  
Db 1510 TCTTTGACTTATTAAGAAAGAAATG-AGTGGGAGTCAAGGCTCTTCTGTGATCTTCTCTA 1568  
Qy 1405 AGAAGAGATTAATTAAGACTGTGTCACTCAACTAAACATTTCTAATATTTTGTAAACATT 1464  
Db 1569 GGAGCAGTTACTTCCAGACTGTGTCACTCAACTAAACCTTCTAATATTTTGTAAACATT 1628  
Qy 1465 AAATTTTGTGACAGTTAAGTAAATATTTGATGTTTGTATGATCAATAGCATATTAATTAAC 1524  
Db 1629 AAATTTTGTGACAGTTAAGTAAATATCTGTATGTTTGTATCAATAGCATATTAATTAAC 1688  
Qy 1525 TTGTTAAGCAATGTTCTTGAATATGATGCAATGAAGAAATTAATAATTTTCTTTT- 1583  
Db 1689 TTGTTAAGTATGTTGTTGATATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1748  
Qy 1584 -TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1642  
Db 1749 ATGTTAATGCACTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCT 1808  
Qy 1643 TGATCTTGGCCCTTTGTACATGAGGCTACCTCTGAAGTGAATTTTTTTT 1690  
Db 1809 TGGTCCCATCTTTTGTACATGAAGGTTGACTCTGAAGTGAATTTTTTTT 1856

RESULT 11  
AAC59283  
ID AAC59283 standard; cDNA; 1222 BP.  
XX  
AC AAC59283;  
XX 02-FEB-2001 (first entry)  
XX

DE Human secreted protein cDNA #7.  
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein; ss.  
 XX  
 OS Homo sapiens.  
 PN WO200056753-A1.  
 XX  
 PD 28-SEP-2000.  
 XX  
 PF 16-MAR-2000; 2000WO-US006765.  
 XX  
 PR 23-MAR-1999; 99US-0126051P.  
 PR 10-DEC-1999; 99US-0169906P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX  
 DR WPI; 2000-594570/56.  
 DR P-PSDB; AAB33724.  
 XX  
 PT Nucleic acid molecules encoding human secreted proteins, used in  
 PT preventing, treating or ameliorating a disorder.  
 XX  
 PS Claim 1; Page 349-350; 410pp; English.  
 XX  
 CC The invention relates to the isolation of genes AAC59277-C59325 encoding  
 CC 49 human secreted proteins AAB33718-B33764. The genes can be used to  
 CC generate fusion proteins by linking to the gene for the human  
 CC immunoglobulin G Fc portion (SEQID1) for increasing the stability of the  
 CC fusion protein as compared to the human protein only. The genes and  
 CC proteins are useful for preventing, ameliorating or treating medical  
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from  
 CC a range of human tissues disclosed in the specification. The nucleic  
 CC acids, proteins, antibodies and (ant)agonists are useful in the  
 CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)  
 CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune rheumatoid arthritis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections  
 XX  
 SQ Sequence 1222 BP; 407 A; 194 C; 229 G; 392 T; 0 U; 0 Other;  
 Query Match 50.3%; Score 1185; DB 3; Length 1222;  
 Best Local Similarity 99.9%; Pred. No. 6.9e-224;  
 Matches 1196; Conservative 0; Mismatches 1; Gaps 1;  
 QY- 517 GTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTACTATCCCAATA 576  
 DB 9 GTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTACTATCCCAATA 68  
 QY 577 GTGTCTCCGATGTGTCAGAGATGCTAGAAATGGTTTGATCATCGTGTGATGTTGTTGTTG 636  
 DB 69 GTGTCTCCGATGTGTCAGAGATGCTAGAAATGGTTTGATCATCGTGTGATGTTGTTGTTG 128  
 QY 637 TGTGTTGAATCTAGGAGCTTACTTACGATTTTCAATTAAGAAAAACAGCTTCTGCCAT 696  
 DB 129 TGTGTTGAATCTAGGAGCTTACTTACGATTTTCAATTAAGAAAAACAGCTTCTGCCAT 188  
 QY 697 TTTCAATTTGACCATCAGGAGATGGGATGATCATCGTGTGATGTTGTTGTTGTTGTTGTTG 756  
 DB 189 TTTCAATTTGACCATCAGGAGATGGGATGATCATCGTGTGATGTTGTTGTTGTTGTTGTTG 248

QY 757 ATCATATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTTGTGTGAAGTCTG 816  
 DB 249 ATCATATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTTGTGTGAAGTCTG 308  
 QY 817 ACTATGTAGTCAAAATATAATTTCTAAATGAAACGTGATGAACGACACACTGAAAAACACAG 876  
 DB 309 ACTATGTAGTCAAAATATAATTTCTAAATGAAACGTGATGAACGACACACTGAAAAACACAG 368  
 QY 877 ATATCAAAAGTTGTTGACTTTTGGAAAGTCAACGATGATGATGAACATCACAGTACTTTGG 936  
 DB 369 ATATCAAAAGTTGTTGACTTTTGGAAAGTCAACGATGATGATGAACATCACAGTACTTTGG 428  
 QY 937 TGTCTACCCGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTT 996  
 DB 429 TGTCTACCCGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTT 488  
 QY 997 GTGATGTTTGGAGCATAGTTGTCATTTCTATTGAAATATTACCTTGGTTTTCACAGTCTTTC 1056  
 DB 489 GTGATGTTTGGAGCATAGTTGTCATTTCTATTGAAATATTACCTTGGTTTTCACAGTCTTTC 548  
 QY 1057 AGACTCATGATAGTAAAGAGCAGCTTGGCAATGATGGAACGAATATTAGGACCCATACAC 1116  
 DB 549 AGACTCATGATAGTAAAGAGCAGCTTGGCAATGATGGAACGAATATTAGGACCCATACAC 608  
 QY 1117 AACACATGATTCAGAAAAACAAAGAAAAACGCAAGTATTTTCCACATAACAGCAGTAGATTGG 1176  
 DB 609 AACACATGATTCAGAAAAACAAAGAAAAACGCAAGTATTTTCCACATAACAGCAGTAGATTGG 668  
 QY 1177 ATGAACACAGTTCTGCTGGTAGATATGTTAGGAGCGCTGCAAAACCGTTGAAGGAATTTA 1236  
 DB 669 ATGAACACAGTTCTGCTGGTAGATATGTTAGGAGCGCTGCAAAACCGTTGAAGGAATTTA 728  
 QY 1237 TGTCTTGTCTGATGATGAAGAACATGAGAAACCTGTTTGGACCTGTTTGAAGAAATGTTAGAA 1296  
 DB 729 TGTCTTGTCTGATGATGAAGAACATGAGAAACCTGTTTGGACCTGTTTGAAGAAATGTTAGAA 788  
 QY 1297 ATGATCCAACTCAAAAGAAATTTACCTTGGATGAAGCAATGCGACATCTCTTTTGTGACTTAT 1356  
 DB 789 ATGATCCAACTCAAAAGAAATTTACCTTGGATGAAGCAATGCGACATCTCTTTTGTGACTTAT 848  
 QY 1357 TAAAAAGAAATGAATGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGATTTAC 1416  
 DB 849 TAAAAAGAAATGAATGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGATTTAC 908  
 QY 1417 TTAAGACTGTGTCAGTCAACTTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGT 1476  
 DB 909 TTAAGACTGTGTCAGTCAACTTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGT 968  
 QY 1477 ACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAG 1536  
 DB 969 ACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAG 1028  
 QY 1537 TATGCTCTTGTATGATGATTAAGAAATTAATAATTTTCTTTTGAATTAACCAT 1596  
 DB 1029 TATGCTCTTGTATGATGATTAAGAAATTAATAATTTTCTTTTGAATTAACCAT 1088  
 QY 1597 TTTAAATACCTTTCAAAATAT -CCTTTTGTGTCAGTGAATAATGATGATGATGATGATGATGAT 1655  
 DB 1089 TTTAAATACCTTTGAATATATACCTTTGTGTCAGTGAATAATGATGATGATGATGATGATGAT 1148  
 QY 1656 TGTACATGGAGGTCACTCTGAAAGTATTTTTTTTGTAGTAAAGAAATCTTGACTA 1712  
 DB 1149 TGTACATGGAGGTCACTCTGAAAGTATTTTTTTTGTAGTAAAGAAATCTTGACTA 1205  
 RESULT 12  
 ADI31373  
 ID ADI31373 standard; cDNA; 1456 BP.  
 XX  
 AC ADI31373;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX



[illegible]







```
Db 1061 GGTCTCAGCCTTGATGTTTGGAGCATAGCTGCATCTTATTGAGTACTACCTTGGT 1120
Qy 1045 TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAG 1104
Db 1121 TCACAGTCTTTTCAGACCCACGATAGTAAAGAGCACCTGGCAATGATGAGCGGATCTTAG 1180
Qy 1105 GACCCATACCAACACACATGATTCAGAAACAGAAACGCAAGTATTTTCCACCAATACC 1164
Db 1181 GACCCATCCAGCAGCATATGATCCAGAAACAGAAACGCAAGTATTTTCCACCAATACC 1240
Qy 1165 AGCTAGATTGGGATGAACACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGCAAAACCGT 1224
Db 1241 AGCTAGATTGGGAGCAGCATATGTTAGCTGGAGATATGTTAGGAGACGCTGCAAGCGT 1300
Qy 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAATCTTTGACCTGTTGCA 1284
Db 1301 TAAAGGAATTTATGCTGTGTCATGACGAAGAGCATGAGAAGCTGTTTGCACCTGGTTCGA 1360
Qy 1285 GAATGTTAGATATGATCAACTCAAGAAATACCTTGGATGAAGCATTCGAGCATCTCTT 1344
Db 1361 GAATGTTGGAGTATGACCCAGGAGAGGATCACCTTGGATGAAGCATTCGAGCACCCCTT 1420
Qy 1345 TCTTTGACTTATTAAAAAGAAATGA 1370
Db 1421 TCTTTGACTTATTAAAAAGAAATGA 1446

RESULT 15
AAD32039 standard; DNA; 21234 BP.
XX AC AAD32039;
XX DT 18-JUN-2002 (first entry)
XX DE Human kinase protein gene.
XX KW Human; kinase protein; enzyme; cytostatic; osteoplastic; gene expression;
KW colon-moderately differentiated adenocarcinoma; chromosome mapping; gene;
KW drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;
KW bone osteosarcoma; single nucleotide polymorphism; SNP; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FH variation replace(75..76,A)
FT FT /*tag= a
FT FT /standard_name= "single nucleotide polymorphism"
FT FT 2007..2059
FT FT /*tag= b
FT FT 2060..3118
FT FT /*tag= c
FT FT 3119..3341
FT FT /*tag= d
FT FT 3342..4462
FT FT /*tag= e
FT FT 4463..4553
FT FT /*tag= f
FT FT 4554..4948
FT FT /*tag= g
FT FT 4949..5015
FT FT /*tag= h
FT FT 5016..8054
FT FT /*tag= i
FT FT replace(7980,C)
FT FT /*tag= j
FT FT /standard_name= "single nucleotide polymorphism"
FT FT 8055..8171
FT FT /*tag= k
FT FT 8172..8258
FT FT /*tag= l
FT FT 8259..8425
FT FT /*tag= m
```

```
FT intron 8426..9007
FT FT /*tag= n
FT FT replace(8571,T)
FT FT /*tag= o
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 9008..9102
FT FT /*tag= p
FT FT 9103..9352
FT FT /*tag= q
FT FT 9353..9482
FT FT /*tag= r
FT FT 9483..13437
FT FT /*tag= s
FT FT replace(11257,Y)
FT FT /*tag= t
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(11684,C)
FT FT /*tag= u
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(13312,C)
FT FT /*tag= v
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 13438..13520
FT FT /*tag= w
FT FT 13521..16514
FT FT /*tag= x
FT FT 16515..16594
FT FT /*tag= y
FT FT 16595..17962
FT FT /*tag= z
FT FT replace(17110,T)
FT FT /*tag= aa
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(17451,C)
FT FT /*tag= ab
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 17963..18053
FT FT /*tag= ac
FT FT 18054..18139
FT FT /*tag= ad
FT FT 18140..18280
FT FT /*tag= ae
FT FT replace(20766,G)
FT FT /*tag= af
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(20913..20914,T)
FT FT /*tag= ag
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT WO200216567-A2.
FT FT 28-FEB-2002.
FT FT 24-AUG-2001; 2001WO-US026389.
FT FT 24-AUG-2000; 2000US-0227470P.
FT FT 19-MAR-2001; 2001US-00810671.
FT FT (APPL-) APPLERA CORP.
FT FT Yan C, Ye J, Ketchum KA, Di Francesco V, Beasley EM;
FT FT WPI; 2002-269354/31.
FT FT New human kinase proteins and nucleic acids, useful in drug screening
FT FT assays, identifying modulators of kinase activity or treating disorders
FT FT characterized by absence or unwanted expression of the protein.
FT FT Claim 4; Fig 3; 81pp; English.
FT FT The invention relates to isolated human kinase proteins and nucleic
FT FT acids. The nucleic acid and peptide sequences can be used as models for
FT FT the development of human therapeutic targets, aid in the identification
```

of therapeutic proteins and serve as targets for the development of human  
therapeutic agents that modulate kinase activity in cells and tissues  
that express the kinase. The nucleic acids are useful as probes or  
primers, in constructing recombinant vectors, for expressing antigenic  
portions of the proteins, chromosome mapping, drug screenings, testing an  
individual for a genotype, and for gene therapy in patients containing  
cells that are aberrant in kinase gene expression. The proteins may be  
used in drug screening assays, in the identification of compounds that  
modulate, stimulate or inhibit kinase activity, in pharmacogenomic  
analysis, in treating disorders characterised by an absence or unwanted  
expression of the protein (bone osteosarcoma, or colon-moderately  
differentiated adenocarcinoma), and in generating antibodies specific for  
the peptides. Such antibodies can be used to detect the protein in situ,  
in vitro, or in cell lysate or supernatant, to isolate and purify the  
proteins from host cells, pharmacogenomic analysis, tissue typing, and in  
inhibiting protein function. The present sequence is human kinase protein  
gene located on chromosome 5

Sequence 21234 BP; 6065 A; 3772 C; 4258 G; 7139 T; 0 U; 0 Other;  
Query Match 46.6%; Score 1097; DB 6; Length 21234;  
Best Local Similarity 100.0%; Pred. No. 2.6e-206;  
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1228 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTGACCTGGTTGGAAGAA 1287  
Db 18138 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTGACCTGGTTGGAAGAA 18197

Qy 1288 TGTTAGAAATGATGCTCAAAAGAAATACCTTGATGGAAGCAATGCGAGCATCCTTTCT 1347  
Db 18198 TGTTAGAAATGATGCTCAAAAGAAATACCTTGATGGAAGCAATGCGAGCATCCTTTCT 18257

Qy 1348 TTGACTTATTAAGAAAGAAATGAATGGAAATCAGTGGTCTTACTATATCTTCTCTAGA 1407  
Db 18258 TTGACTTATTAAGAAAGAAATGAATGGAAATCAGTGGTCTTACTATATCTTCTCTAGA 18317

Qy 1408 AGAGATTACTTTAAGACTGTGTCAGTCAACTAAACATTTCTAATTTTGTAAACATTTAAA 1467  
Db 18318 AGAGATTACTTTAAGACTGTGTCAGTCAACTAAACATTTCTAATTTTGTAAACATTTAAA 18377

Qy 1468 TTATTTGTACAGTTAAGTGAATAATATGTAATGTTTGTATCAATAGCATAAATTAACCTTG 1527  
Db 18378 TTATTTGTACAGTTAAGTGAATAATATGTAATGTTTGTATCAATAGCATAAATTAACCTTG 18437

Qy 1528 TTAAGCAAGTATGCTTTGTAATGCAATTAGAAAATTTAAATTTCTTTCTTTTGAA 1587  
Db 18438 TTAAGCAAGTATGCTTTGTAATGCAATTAGAAAATTTAAATTTCTTTCTTTTGAA 18497

Qy 1588 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGCCAGTGATAAATGTAATGATC 1647  
Db 18498 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGCCAGTGATAAATGTAATGATC 18557

Qy 1648 TTGCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTTTTTGTAGTAAAGGAAATCTT 1707  
Db 18558 TTGCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTTTTTGTAGTAAAGGAAATCTT 18617

Qy 1708 GACTACTTTTATTTCTTAAAGGAAATTTCTTTTATATCTTCAAAATTTAGAACTTTAACTTT 1767  
Db 18618 GACTACTTTTATTTCTTAAAGGAAATTTCTTTTATATCTTCAAAATTTAGAACTTTAACTTT 18677

Qy 1768 AAAAGTTTTTCTCTCTGTAATTTGTTGAACGGGTGATTAATTAATTAACCTPAGATAAGCAGGT 1827  
Db 18678 AAAAGTTTTTCTCTCTGTAATTTGTTGAACGGGTGATTAATTAATTAACCTPAGATAAGCAGGT 18737

Qy 1828 ACTAGAAACCAAACTCAGAAAAATGTTTACGTGTAAGAAATCTTATTAATTTTAAGTGTG 1887  
Db 18738 ACTAGAAACCAAACTCAGAAAAATGTTTACGTGTAAGAAATCTTATTAATTTTAAGTGTG 18797

Qy 1888 TATTCTTTTTCATTTGGGTGATGTCAGGGTGAATAACAGACATTCATGGAAGGCATGCAG 1947  
Db 18798 TATTCTTTTTCATTTGGGTGATGTCAGGGTGAATAACAGACATTCATGGAAGGCATGCAG 18857

Qy 1948 TTTGTCCATTGTGACAGTTTGTGTTTAAATAAAACCATACATACACACTTTTATTTAAGATTAAA 2007

Db 18858 TTTGTCCATTGTGACAGTTTGTGTTTAAATAAAACCATACACACTTTTATTTAAGATTAAAA 18917  
Qy 2008 TCTTAACCTGGAAGTCAGCTTGGAAAATGGACATTTCCAAAGTATGTTTGGTGCATCACAGA 2067  
Db 18918 TCTTAACCTGGAAGTCAGCTTGGAAAATGGACATTTCCAAAGTATGTTTGGTGCATCACAGA 18977  
Qy 2068 TATAAAATAGAAAATTTCTGATGAGAGGTTTTCAGATTTTAAATACCAAGTCCTTTAGGAGTCT 2127  
Db 18978 TATAAAATAGAAAATTTCTGATGAGAGGTTTTCAGATTTTAAATACCAAGTCCTTTAGGAGTCT 19037  
Qy 2128 TAAATTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTTATAGAAATTAAG 2187  
Db 19038 TAAATTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTTATAGAAATTAAG 19097  
Qy 2188 TTTTATTAATTAGGCAATTTATGCTCTGTGATTAATTTCTTACGGGAGAAAGAGATTTTGATTG 2247  
Db 19098 TTTTATTAATTAGGCAATTTATGCTCTGTGATTAATTTCTTACGGGAGAAAGAGATTTTGATTG 19157

Qy 2248 GAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAAATGATTGGTTACAT 2307  
Db 19158 GAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAAATGATTGGTTACAT 19217

Qy 2308 AAATTTTGTACTTTCAG 2324  
Db 19218 AAATTTTGTACTTTCAG 19234

Search completed: March 16, 2005, 09:01:13  
Job time : 1257.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 07:34:52 ; Search time 7615.01 Seconds  
(without alignments)  
11766.665 Million cell updates/sec

Title: US-10-801-671-1  
Perfect score: 2354  
Sequence: 1 gccagctgggttacttta.....taaaaaaacaaaaaaac 2354

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hic:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_ges1:  
9: gb\_ges2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID        | Description         |
|------------|--------|-------------|--------|----|-----------|---------------------|
| 1          | 1689.4 | 71.8        | 2798   | 3  | CR749504  | CR749504 Homo sapi  |
| 2          | 1283.2 | 54.5        | 1556   | 3  | BC015942  | BC015942 Homo sapi  |
| 3          | 1194   | 50.7        | 1446   | 9  | AY408252  | AY408252 Homo sapi  |
| 4          | 1118.8 | 47.5        | 1865   | 3  | AF212224  | AF212224 Homo sapi  |
| 5          | 1078.2 | 45.8        | 1629   | 3  | AK013974  | AK013974 Mus muscu  |
| 6          | 978.2  | 41.6        | 1421   | 9  | AY408253  | AY408253 Mus muscu  |
| 7          | 937.6  | 39.8        | 1446   | 9  | AY408253  | AY408253 Pan trogl  |
| 8          | 814.8  | 34.6        | 1737   | 3  | CR591789  | CR591789 full-leng  |
| 9          | 814.8  | 34.6        | 1746   | 3  | CR612115  | CR612115 full-leng  |
| 10         | 814.8  | 34.6        | 1754   | 3  | CR617472  | CR617472 full-leng  |
| 11         | 754.4  | 32.0        | 973    | 5  | BUS16568  | BUS16568 AGENCOURT  |
| 12         | 734.2  | 31.2        | 1689   | 3  | AK076199  | AK076199 Mus muscu  |
| 13         | 722    | 30.7        | 722    | 4  | BM786882  | BM786882 K-EST0065  |
| 14         | 720.8  | 30.6        | 966    | 7  | CF110608  | CF110608 Shultzomi  |
| 15         | 700.8  | 29.8        | 727    | 1  | AA631990  | AA631990 np74a05.s  |
| 16         | 700.6  | 29.8        | 729    | 4  | BM675318  | BM675318 UI-EJ0-    |
| 17         | 670.6  | 28.5        | 695    | 5  | BQ774390  | BQ774390 UI-H-EZ1-  |
| 18         | 666    | 28.3        | 1664   | 3  | CR593610  | CR593610 full-leng  |
| 19         | 647.6  | 27.5        | 1775   | 3  | BC028149  | BC028149 Homo sapi  |
| 20         | 644.8  | 27.4        | 721    | 5  | BUI184090 | BUI184090 AGENCOURT |
| 21         | 643.2  | 27.3        | 659    | 4  | BM681648  | BM681648 UI-EJ0-    |
| 22         | 641.4  | 27.2        | 664    | 5  | BM969890  | BM969890 UI-CF-DU1  |
| 23         | 640    | 27.2        | 3226   | 3  | BC028573  | BC028573 Homo sapi  |
| 24         | 635.2  | 27.0        | 643    | 6  | CB127989  | CB127989 K-EST0177  |

|    |       |      |      |   |          |                    |
|----|-------|------|------|---|----------|--------------------|
| 25 | 635.2 | 27.0 | 919  | 4 | BG109862 | BG109862 602279528 |
| 26 | 632.2 | 26.9 | 637  | 5 | BUT29717 | BUT29717 UI-E-CK1- |
| 27 | 628   | 26.7 | 1337 | 7 | CO389096 | CO389096 AGENCOURT |
| 28 | 621.8 | 26.4 | 849  | 4 | BG916444 | BG916444 602814061 |
| 29 | 620.6 | 26.4 | 667  | 1 | AJ656344 | AJ656344 AJ656344  |
| 30 | 617   | 26.2 | 1551 | 7 | CF110517 | CF110517 Shultzomi |
| 31 | 611.6 | 26.0 | 957  | 5 | BX374503 | BX374503 BX374503  |
| 32 | 609.4 | 25.9 | 923  | 5 | BX385193 | BX385193 BX385193  |
| 33 | 609   | 25.9 | 767  | 5 | BX374502 | BX374502 BX374502  |
| 34 | 607   | 25.8 | 607  | 6 | CD676847 | CD676847 ho07d04.y |
| 35 | 594.8 | 25.3 | 1138 | 3 | CR609176 | CR609176 full-leng |
| 36 | 592.2 | 25.2 | 902  | 1 | AL540351 | AL540351 AL540351  |
| 37 | 591.6 | 25.1 | 950  | 7 | CF411264 | CF411264 CH3#072.F |
| 38 | 591   | 25.1 | 624  | 4 | BI494554 | BI494554 df11a06.  |
| 39 | 589.8 | 25.1 | 593  | 6 | CA314357 | CA314357 UI-CF-FNO |
| 40 | 589.8 | 25.1 | 893  | 6 | CB202758 | CB202758 AGENCOURT |
| 41 | 589.2 | 25.0 | 950  | 5 | BP383242 | BP383242 BP383242  |
| 42 | 573.4 | 24.4 | 710  | 5 | BQ210018 | BQ210018 UI-R-EPO- |
| 43 | 570.8 | 24.2 | 618  | 6 | CB285812 | CB285812 CMD25.A05 |
| 44 | 570   | 24.2 | 581  | 5 | BP307701 | BP307701 BP307701  |
| 45 | 566   | 24.0 | 568  | 4 | BM727179 | BM727179 UI-E-EJ0- |

## ALIGNMENTS

RESULT 1  
LOCUS CR749504 2798 bp mRNA linear HTC 19-AUG-2004  
DEFINITION Homo sapiens mRNA; cDNA DKFP686A20267 (from clone DKFP686A20267).  
ACCESSION CR749504  
VERSION CR749504.1 GI:51476700  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2798)  
AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,  
Fobo, G., Han, M. and Wiemann, S.  
CONSRMT The German cDNA Consortium  
TITLE Direct Submission  
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by BMPZ (Biomedical Research Center at the  
Heinrich-Heine-University, Dueseldorff/Germany) within the cDNA  
sequencing consortium of the German Genome Project. This clone  
(DKFP686A20267) is available at the RZPD Deutsches  
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFP686A20267  
Further information about the clone and the sequencing project is  
available at http://mips.gsf.de/projects/cdna/.

## FEATURES

source  
1..2798  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFP686A20267"  
/tissue type="cerebellum"  
/clone\_lib="686 (synonym: hlccc3). Vector pSport1\_sfi; host  
DH10B; sites SfiI + SfiIIB"  
/dev stage="adult"  
/note="protein serine threonine kinase Ctk4,  
differentially spliced"  
1..2798  
/gene="DKFP686A20267"  
85..246  
/gene="DKFP686A20267"  
/codon\_start=1  
/product="hypothetical protein"

/protein\_id="CAH18325.1"  
/db\_xref="GI:51476701"  
/translation="WRHSKRTHCPDWDSDRESWGHSYRGSHKRRKSHSSSQENRHCK  
PHHQPKESD"

ORIGIN

|                       |        |              |         |            |     |        |       |
|-----------------------|--------|--------------|---------|------------|-----|--------|-------|
| Query Match           | 71.8%; | Score        | 1689.4; | DB         | 3;  | Length | 2798; |
| Best Local Similarity | 99.4%; | Pred. No.    | 0;      | Mismatches | 11; | Indels | 0;    |
| Matches               | 1696;  | Conservative | 0;      | Gaps       | 0;  |        |       |

  

|    |      |   |      |
|----|------|---|------|
| QY | 27   | TGCTCCATGTGATCCCTCTTGAAGCTTCGACCTCTGTTGAAGAGGACATCATCCAGT   | 86   |
| DB | 1090 | TGCTCCATGTGATCCCTCTTGAAGCTTCGACCTCTGTTGAAGAGGACATCATCCAGT   | 1149 |
| QY | 87   | CATTATTTAGAGCAAGCTCTTGAATCAGCGAGATTATCGGACCGAGATACCTTCAC    | 146  |
| DB | 1150 | CATTATTTAGAGCAAGCTCTTGAATCAGCGAGATTATCGGACCGAGATACCTTCAC    | 1209 |
| QY | 147  | GAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTATCACAGAGACATTGAA | 206  |
| DB | 1210 | GAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTATCACAGAGACATTGAA | 1269 |
| QY | 207  | AGCGGATTCGAATCCACTCGAGTAATCTTCAGTCCGACGAGAGAGAGAGAGAGAGAG   | 266  |
| DB | 1270 | AGCGGATTCGAATCCACTCGAGTAATCTTCAGTCCGACGAGAGAGAGAGAGAGAGAG   | 1329 |
| QY | 267  | AGGAGCGCAATAGACACTGTTCAAGTCATCAGTCACTGTCGAGAGCCACCGAAGGAA   | 326  |
| DB | 1330 | AGGAGCGCAATAGACACTGTTCAAGTCATCAGTCACTGTCGAGAGCCACCGAAGGAA   | 1389 |
| QY | 327  | AGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACGTT  | 386  |
| DB | 1390 | AGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACGTT  | 1449 |
| QY | 387  | CTAAGAGCAAGATGAAATCGTGAACACTTGGGTGAAGAGCCCTTTGGCAAGTTGTA    | 446  |
| DB | 1450 | CTAAGAGCAAGATGAAATCGTGAACACTTGGGTGAAGAGCCCTTTGGCAAGTTGTA    | 1509 |
| QY | 447  | GAGTCATTGATGATGAGGATGATGAGGATGATGAGGATGATGAGGATGATGAGGATGAT | 506  |
| DB | 1510 | GAGTCATTGATGATGAGGATGATGAGGATGATGAGGATGATGAGGATGATGAGGATGAT | 1569 |
| QY | 507  | GSCCGTTACCGTGAAGAGCTCGTTTCAAGATCCAAATTTAGAGCACTTAAATAGTACT  | 566  |
| DB | 1570 | GSCCGTTACCGTGAAGAGCTCGTTTCAAGATCCAAATTTAGAGCACTTAAATAGTACT  | 1629 |
| QY | 567  | GATCCCAATAGTCTTCGATGTCGAGATGTCGAGATGTCGAGATGTCGAGATGTCGAG   | 626  |
| DB | 1630 | GATCCCAATAGTCTTCGATGTCGAGATGTCGAGATGTCGAGATGTCGAGATGTCGAG   | 1689 |
| QY | 627  | GTTCGTATTTGTTGAATCTCTGGACTTACTACTTACGATTTTCAATTAAGAAACACAGC | 686  |
| DB | 1690 | GTTCGTATTTGTTGAATCTCTGGACTTACTACTTACGATTTTCAATTAAGAAACACAGC | 1749 |
| QY | 687  | TTTCTGCCATTTCAAATTTGACCATCAGGAGATGGCGTATCAGATCTGCCAGTCAATA  | 746  |
| DB | 1750 | TTTCTGCCATTTCAAATTTGACCATCAGGAGATGGCGTATCAGATCTGCCAGTCAATA  | 1809 |
| QY | 747  | AAATTTTACATCAATAAATTAACCCATCAGATCTGAAGCCCTGAAATATTTTGGTT    | 806  |
| DB | 1810 | AAATTTTACATCAATAAATTAACCCATCAGATCTGAAGCCCTGAAATATTTTGGTT    | 1869 |
| QY | 807  | GTGAAGTCTGACTATGTAGTCAATATAATTTCTAAATGAAACGATGAGCAACACTG    | 866  |
| DB | 1870 | GTGAAGTCTGACTATGTAGTCAATATAATTTCTAAATGAAACGATGAGCAACACTG    | 1929 |
| QY | 867  | AAAAACACAGATATCAAAGTTGTTGACCTTGAAGTGAACCGATGATGATGAACATCAC  | 926  |
| DB | 1930 | AAAAACACAGATATCAAAGTTGTTGACCTTGAAGTGAACCGATGATGATGAACATCAC  | 1989 |
| QY | 927  | AGTACTTTGGTCTTACCGGCACTACAGAGCTCCCGAGTCAATTTGGCTTTAGTTGG    | 986  |
| DB | 1990 | AGTACTTTGGTCTTACCGGCACTACAGAGCTCCCGAGTCAATTTGGCTTTAGTTGG    | 2049 |

|    |      |  |      |
|----|------|--|------|
| QY | 987  | TCTCAGCCTTGTGTATGTTTGGAGCATAGGTTGCATCTTATTGAATATTACCTTGGTTTC   | 1046 |
| DB | 2050 | TCTCAGCCTTGTGTATGTTTGGAGCATAGGTTGCATCTTATTGAATATTACCTTGGTTTC   | 2109 |
| QY | 1047 | ACAGTCTTTTCAAGCTCATGATAGTAAAGAGACACCTGGCAATGATGGAACGATATTAGGA  | 1106 |
| DB | 2110 | ACAGTCTTTTCAAGCTCATGATAGTAAAGAGACACCTGGCAATGATGGAACGATATTAGGA  | 2169 |
| QY | 1107 | CCCATACCAACACATGATTTCAGAAAACAGAAAACGCAAGTATTTTTCACATAACACAG    | 1166 |
| DB | 2170 | CCCATACCAACACATGATTTCAGAAAACAGAAAACGCAAGTATTTTTCACATAACACAG    | 2229 |
| QY | 1167 | CTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAACCCGTTG  | 1226 |
| DB | 2230 | CTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAACCCGTTG  | 2289 |
| QY | 1227 | AAGGAATTTATGCTTTGTCTGATGATGAAGACATGAGAACTGTTTGGCTGGTTCGAGA     | 1286 |
| DB | 2290 | AAGGAATTTATGCTTTGTCTGATGATGAAGACATGAGAACTGTTTGGCTGGTTCGAGA     | 2349 |
| QY | 1287 | ATGTTAGAAATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCAATTCAGCATCTTTC | 1346 |
| DB | 2350 | ATGTTAGAAATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCAATTCAGCATCTTTC | 2409 |
| QY | 1347 | TTTGACTTATTAAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTCTCTAG   | 1406 |
| DB | 2410 | TTTGACTTATTAAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTCTCTAG   | 2469 |
| QY | 1407 | AAGAGATTACTTAAGACCTGTCAGTCACTAAACATTTCTAATATTTTGTAAACATTA      | 1466 |
| DB | 2470 | AAGAGATTACTTAAGACCTGTCAGTCACTAAACATTTCTAATATTTTGTAAACATTA      | 2529 |
| QY | 1467 | ATTATTTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATAATTAACCTT     | 1526 |
| DB | 2530 | ATTATTTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATAATTAACCTT     | 2589 |
| QY | 1527 | GTTAAGCAAGTATGCTTTGATTAATGCAATGAGAAATTAATAATTTTCTTTTGA         | 1586 |
| DB | 2590 | GTTAAGCAAGTATGCTTTGATTAATGCAATGAGAAATTAATAATTTTCTTTTGA         | 2649 |
| QY | 1587 | AATTACCAATTTTAAATACCTTTGAATATCTTTTGTCCAGTCATGATGATGATGAT       | 1646 |
| DB | 2650 | AATTACCAATTTTAAATACCTTTGAATATCTTTTGTCCAGTCATGATGATGATGAT       | 2709 |
| QY | 1647 | CTTGCTTTTGTACATGAGGCTCACCTCTGAAGTGAATTTTGTGAGTAAAGGAAATCT      | 1706 |
| DB | 2710 | CTTGCTTTTGTACATGAGGCTCACCTCTGAAGTGAATTTTGTGAGTAAAGGAAATCT      | 2769 |
| QY | 1707 | TGACTACTTTTATTTCTTTAAAGGAATA                                   | 1733 |
| DB | 2770 | TGACTACTTTTAAAGGAAATTTTAAAGGAAATA                              | 2796 |

|            |   |
|------------|---|
| RESULT 2   |   |
| BC015942   |   |
| LOCUS      | 1556 bp mRNA linear HTC 19-NOV-2003   |
| DEFINITION | Homo sapiens CDC-like kinase 4, mRNA (cdna clone IMAGE:3908500), containing frame-shift errors.   |
| ACCESSION  | BC015942.1 GI:16198521  |
| VERSION    | HTC.  |
| KEYWORDS   | Homo sapiens (human)  |
| SOURCE     | Homo sapiens  |
| ORGANISM   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.   |
| REFERENCE  | 1 (bases 1 to 1556)   |
| AUTHORS    | Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., |

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257  
12477932  
2 (bases 1 to 1556)  
Straussberg, R.  
Direct Submission  
Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 15 Row: d Column: 16  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10190705  
This clone has the following problem: frame shifted.

## FEATURES

Location/Qualifiers  
1..1556  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3908500"  
/tissue\_type="Uterus, leiomyosarcoma"  
/clone\_lib="NIH MGC\_71"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

## ORIGIN

Query Match 54.5%; Score 1283.2; DB 3; Length 1556;  
Best Local Similarity 99.8%; Pred. No. 2.2e-254;  
Matches 1285; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 85 GTCATTATTAGAACGAGGTCCTTGAATGACGAGATATTCGGACCGGAGATACGTTG 144  
DB 269 GTCATTATTAGAACGAGGTCCTTGAATGACGAGATATTCGGACCGGAGATACGTTG 328

QY 145 ACGAATACAGGAACTGACTCTGTAAGGATATGTTCCCTAGACATTATCACAGACATTG 204  
DB 329 ACGAATACAGGAACTGACTCTGTAAGGATATGTTCCCTAGACATTATCACAGACATTG 388

QY 205 AAAGCGGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGACGAGGAGACGTCCTTA 264

DB 389 AAAGCGGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGACGAGGAGACGTCCTTA 448  
QY 265 AAAGGAGCGCAATAGACACCTGTTTCAAGTCATCAGTCACGTTTCAAGGAGCCACCGAAGGA 324  
DB 449 AAAGGAGCGCAATAGACACCTGTTTCAAGTCATCAGTCACGTTTCAAGGAGCCACCGAAGGA 508  
QY 325 AAAGTCCAGGAGTATAGAGATAGATAGAGAGGTCACCTGATCTGTCAAAGTGAGAGCG 384  
DB 509 AAAGTCCAGGAGTATAGAGATAGATAGAGAGGTCACCTGATCTGTCAAAGTGAGAGCG 568  
QY 385 TTCTAAGGAGCAAGTATGAAATCGTGGACACCTTTGGGTGAAGGAGCCCTTTGCAAGTTG 444  
DB 569 TTCTAAGGAGCAAGTATGAAATCGTGGACACCTTTGGGTGAAGGAGCCCTTTGCAAGTTG 628  
QY 445 TAGAGTGCATTGATCATCGCATGATGCGCATGTCATGTAGCAGTCAAAATTCGTAAATAATG 504  
DB 629 TAGAGTGCATTGATCATCGCATGATGCGCATGTCATGTAGCAGTCAAAATTCGTAAATAATG 688  
QY 505 TAGGCGGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCCTTAAATAGTA 564  
DB 689 TAGGCGGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCCTTAAATAGTA 748  
QY 565 CTGATCCCAATAGTGTCTTCCGATGTCGTCAGATGCTAGATGTTTTCATCATCATGTC 624  
DB 749 CTGATCCCAATAGTGTCTTCCGATGTCGTCAGATGCTAGATGTTTTCATCATCATGTC 808  
QY 625 ATGTTTGTATGTTGTTTCAACTTCTGGGACCTTAGTACTTACGATTTCAATTAAGAAACA 684  
DB 809 ATGTTTGTATGTTGTTTCAACTTCTGGGACCTTAGTACTTACGATTTCAATTAAGAAACA 868  
QY 685 GCTTTCTGCCATTTCAAATTCGACCATCATGAGCAGATCGCGTATCAGATCTGCCAGTCAA 744  
DB 869 GCTTTCTGCCATTTCAAATTCGACCATCATGAGCAGATCGCGTATCAGATCTGCCAGTCAA 928  
QY 745 TAAATTTTACATCATTAATAATTAACCCATACAGATCTGAGGCTGAGATATTTTGT 804  
DB 929 TAAATTTTACATCATTAATAATTAACCCATACAGATCTGAGGCTGAGATATTTTGT 988  
QY 805 TTGTGAAGTCTGACTATCTAGTCAAAATATAATTTCTTAAATGAAAGCTGATGAACGACAC 864  
DB 989 TTGTGAAGTCTGACTATCTAGTCAAAATATAATTTCTTAAATGAAAGCTGATGAACGACAC 1048  
QY 865 TGAATAACACAGATATCAAGATGTTTGTGTTTGGAGTGCACACGATGATGATGAACATC 924  
DB 1049 TGAATAACACAGATATCAAGATGTTTGTGTTTGGAGTGCACACGATGATGATGAACATC 1108  
QY 925 ACAGTACTTTGTTGTTCTACCCGACATCAGAGTCCCGAGGTCATTTTGGCTTTAGGTT 984  
DB 1109 ACAGTACTTTGTTGTTCTACCCGACATCAGAGTCCCGAGGTCATTTTGGCTTTAGGTT 1168  
QY 985 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTTCATTTTGAATATTTACCTTTGGTT 1044  
DB 1169 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTTCATTTTGAATATTTACCTTTGGTT 1228  
QY 1045 TCAGAGTCTTTTTCAGATCTCATGATAGTAAAGAGCAGCTTGGCAATGATGGAACGAATTTAG 1104  
DB 1229 TCAGAGTCTTTTTCAGATCTCATGATAGTAAAGAGCAGCTTGGCAATGATGGAACGAATTTAG 1288  
QY 1105 GACCCATACCAACACATGATTCGAAAAACAAGAAACGCAAGTATTTTTCACCAATAACC 1164  
DB 1289 GACCCATACCAACACATGATTCGAAAAACAAGAAACGCAAGTATTTTTCACCAATAACC 1348  
QY 1165 AGCTAGATTGGAGTGAACACAGTTCCTGCTGGTAGATATGTTAGGAGACGCTGCAACCGT 1224  
DB 1349 AGCTAGATTGGAGTGAACACAGTTCCTGCTGGTAGATATGTTAGGAGACGCTGCAACCGT 1408  
QY 1225 TGAAGGAATTTTATGTTTGTTCATGATGAAGAACATGAGAAACTGTTTGAACCTGGTTGAA 1284  
DB 1409 TGAAGGAATTTTATGTTTGTTCATGATGAAGAACATGAGAAACTGTTTGAACCTGGTTGAA 1468  
QY 1285 GAATGTTTGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTTGAGCAATCCTT 1344

Db 1469 CAATGTTAGAAATGATCCAACTCAAGCAATTAATCCTTGGATGAAGCAATTCAGCATCCTT 1528  
QY 1345 TCTTTGACTTATTAAGAAAGAAATGAAA 1372  
Db 1529 TCTTTGACTTATTAAGAAAGAAATGAAA 1556

RESULT 3  
AY408252 1446 bp DNA linear GSS 15-DEC-2003  
LOCUS Homo sapiens CLK4 gene, VIRTUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION AY408252  
VERSION AY408252.1 GI:39764223  
KEYWORDS GSS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1446)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
Science 302 (5652), 1960-1963 (2003)  
JOURNAL 14671302  
PUBMED 2 (bases 1 to 1446)  
REFERENCE Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
AUTHORS Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Direct Submission  
TITLE Submitted (18-NOV-2003) Celera Genomics, 45 West Gude Drive,  
JOURNAL Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source 1..1446  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
gene <1..>1446  
/gene="CLK4"  
/locus\_tag="HCM3162"

ORIGIN  
Query Match 50.7%; Score 1194; DB 9; Length 1446;  
Best Local Similarity 92.8%; Pred. No. 5.9e-236;  
Matches 1194; Conservative 0; Mismatches 92; Indels 0; Gaps 0;  
QY 85 GTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 144  
Db 161 GTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 220  
QY 145 ACGAATACAGGAATGACTACTGCTGAAGGATATGTTCTTAGACATTATCACAGACATTG 204  
Db 221 ACGAATACAGGAATGACTACTGCTGAAGGATATGTTCTTAGACATTATCACAGACATTG 280  
QY 205 AAAGCGGTATCGAATCCATGTCAGTAATCTTCAGTCCGACGAGGAGAGAGATGTCCTTA 264  
Db 281 AAAGCGGTATCGAATCCATGTCAGTAATCTTCAGTCCGACGAGGAGAGAGATGTCCTTA 340  
QY 265 AAAGCAAGCGCAATAGACACTGTTCAAGTCATCACTGCTGAGGACCGACCCCAAGGA 324  
Db 341 AAAGCAAGCGCAATAGACACTGTTCAAGTCATCACTGCTGAGGACCGACCCCAAGGA 400  
QY 325 AAAGATCCAGGAGTATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGCG 384  
Db 401 AAAGATCCAGGAGTATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGCG 460  
QY 385 TTCTAAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGCCCTTTGGCAAGTTG 444

Db 461 TTTCTAAGAGCAAGATNN 520  
QY 445 TAGAGTGCATTGATCATGGCATGGCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATG 504  
Db 521 NNN 580  
QY 505 TAGGCCGTTACCGTGAAGCAGCTCGTTTCAAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 564  
Db 581 TAGGCCGTTACCGTGAAGCAGCTCGTTTCAAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 640  
QY 565 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGATGCTAGATGCTAGATGCTAGATG 624  
Db 641 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGATGCTAGATGCTAGATGCTAGATG 700  
QY 625 ATGTTTGTATTGTCTTGAACACTTCTGGGACTTCTAGTCTTCAATTTCAATTAAGAAACA 684  
Db 701 ATGTTTGTATTGTCTTGAACACTTCTGGGACTTCTAGTCTTCAATTTCAATTAAGAAACA 760  
QY 685 GCTTTCTGCCATTTCAAATTTGACCAATCATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 744  
Db 761 GCTTTCTGCCATTTCAAATTTGACCAATCATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 820  
QY 745 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAAGCTGAAATATTTTGT 804  
Db 821 NNNNNNTTTTACATCATATAAATTAACCCATACAGATCTGAAAGCTGAAATATTTTGT 880  
QY 805 TTGTGAAGTCTGACTATGTAGTCAATATATAATTTCTAAATGAAACGCTGATGAAGCACAC 864  
Db 881 TTGTGAAGTCTGACTATGTAGTCAATATATAATTTCTAAATGAAACGCTGATGAAGCACAC 940  
QY 865 TGAAGAAACACAGATATCAAAGTGTGTGCTTTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 924  
Db 941 TGAAGAAACACAGATATCAAAGTGTGTGCTTTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1000  
QY 925 ACAGTACTTTTGTCTTACCCGCACTACAGATCTCCGAGGTCTATTTTGGCTTTAGTT 984  
Db 1001 ACAGTACTTTTGTCTTACCCGCACTACAGATCTCCGAGGTCTATTTTGGCTTTAGTT 1060  
QY 985 GGTCTCAGCCCTTGTGATGTTTGGAGCATAGTGTGCTTCTTATTGAATATTACCTTGGTT 1044  
Db 1061 GGTCTCAGCCCTTGTGATGTTTGGAGCATAGTGTGCTTCTTATTGAATATTACCTTGGTT 1120  
QY 1045 TCAGATCTTTTCACTCATGATAGTAAAGAGACCTGGCAATGATGAAGCAATATTAG 1104  
Db 1121 TCAGATCTTTTCACTCATGATAGTAAAGAGACCTGGCAATGATGAAGCAATATTAG 1180  
QY 1105 GACCCATACCAACACATGATTTTCAAGAAACGCAAGTATTTTTCAACCAATACC 1164  
Db 1181 GACCCATACCAACACATGATTTTCAAGAAACGCAAGTATTTTTCAACCAATACC 1240  
QY 1165 AGCTAGATTGGGATGAACACAGTCTTCTGCTGTAGATATGTTTAGGAGACGCTGCAACCGT 1224  
Db 1241 AGCTAGATTGGGATGAACACAGTCTTCTGCTGTAGATATGTTTAGGAGACGCTGCAACCGT 1300  
QY 1225 TGAAGAAATTTATGCTTTGTGATGAAGAAACATGAGAAACTGTTTGAACCTGTTTGA 1284  
Db 1301 TGAAGAAATTTATGCTTTGTGATGAAGAAACATGAGAAACTGTTTGAACCTGTTTGA 1360  
QY 1285 GAATGTTAGATATCATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTTGAGCATCTTT 1344  
Db 1361 GAATGTTAGATATCATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTTGAGCATCTTT 1420  
QY 1345 TCTTTGACTTATTAAGAAAGAAATGA 1370  
Db 1421 TCTTTGACTTATTAAGAAAGAAATGA 1446

RESULT 4  
AF212224  
LOCUS Homo sapiens CLK4 mRNA, complete cds.  
DEFINITION AF212224  
ACCESSION AF212224

linear HTC 22-MAY-2001  
1865 bp mRNA  
complete cds.

|                            |   |
|----------------------------|---|
| AF212224.1                 | GI:9437514  |
| HTC.                       |   |
| KEYWORDS                   |   |
| SOURCE                     |   |
| ORGANISM                   |   |
| REFERENCE                  |   |
| AUTHORS                    |   |
| TITLE                      |   |
| JOURNAL                    |   |
| REFERENCE                  |   |
| AUTHORS                    |   |
| TITLE                      |   |
| JOURNAL                    |   |
| FEATURES                   |   |
| source                     |   |
| CDS                        |   |
| ORIGIN                     |   |
| Query Match                | 47.5%; Score 1118.8; DB 3; Length 1865;                                 |
| Best Local Similarity      | 90.2%; Pred. No. 2e-220;  |
| Matches 1191; Conservative | 0; Mismatches 129; Indels 0; Gaps 0;                                    |
| Qy                         | 400 ATGAAATCGTGGACACTTTGGGTGAAGGAGCGCTTTGGCAAAGTTGTAGAGTGCATTCGATC 459  |
| Db                         |   |
| Qy                         | 545 ATGAAATGTTGTATACCTTTAGTCAAGGAGCGCTTTGGGAAAAGTTGTGGAGTGCATCGATC 604  |
| Db                         |   |
| Qy                         | 460 ATGCGATGGATGGCATCGCATGTAGCAGTGAATAATCGTAAAAAATGTAGGCGGTTCACGCTG 519 |
| Db                         |   |
| Qy                         | 605 ATAAAGCGGGAGGTAGACATGTAGCAGTGAATAATAGTAAAAATGTGGATAGATACACTGTG 664  |
| Db                         |   |
| Qy                         | 520 AAGCAGCTCGTTTCAGAAATCAAGTATTAGAGCACTTTAAATAGTACTGATCCCAATAGTG 579   |
| Db                         |   |
| Qy                         | 665 AAGCTGCTCGCTCAGAAATACAAGTTCTGGAACATCTGATATACAACAGACCCCAACAGTA 724   |
| Db                         |   |
| Qy                         | 580 TCTTCCGATGTGTCAGATGCTAGAAATGGTTTGATCATCATGTGTCATGTTGTATTGTGTG 639   |
| Db                         |   |
| Qy                         | 725 CTTTCCGCTGTGTCAGATGTTGGAATGGTTTGGAGCATCATGGTCACATTTGCATTGCTTT 784   |
| Db                         |   |
| Qy                         | 640 TTGAACACTCGGGACTTAGTACTTACAGATTTCATTAAGAAAACAGACTTCTTCGCCATTTC 699  |
| Db                         |   |
| Qy                         | 785 TTGAACATTATGGGACTTAGTACTTACGACTTCATTAAGAAAATGGTTTCTTACCATTTTC 844   |
| Db                         |   |
| Qy                         | 700 AAATTGACCACATCAGGCAGATGCGGTATCAGATCTGCCAGTCAATAAATTTTTTACATC 759    |
| Db                         |   |
| Qy                         | 845 GACTGGATCATATCAGAAAGATGCAATATCAGATATGCAAGTCTGTGAAATTTTTTGCACA 904   |
| Db                         |   |
| Qy                         | 760 ATAATAAATAAACCCATCAGATCTGAAGCCTGAAAAATTTTTGTTGTGGAAGTCTGACT 819     |
| Db                         |   |
| Qy                         | 905 GTAATAAGTTGACTTCACACAGACTTAAAGCCTGAAAACATCTTATTTGTGCGACTCTGACT 964  |
| Db                         |   |
| Qy                         | 820 ATGTAAGTCAATAATAATCTTAAATGAAACGTGATGAACGCACACTCGAAAAACACAGATA 879   |
| Db                         |   |
| Qy                         | 965 ACACAGAGGGCGTAAATCCCAATAATAAAAGTGAATGAACGCACTTAATAATCAAGATA 1024    |
| Db                         |   |

|           |  |
|-----------|--|
| TITLE     | High-efficiency full-length cDNA cloning   |
| JOURNAL   | Meth. Enzymol. 303, 19-44 (1999)   |
| MEDLINE   | 99279253   |
| PUBMED    | 10349636   |
| REFERENCE | 2  |
| AUTHORS   | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.   |
| TITLE     | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes   |
| JOURNAL   | Genome Res. 10 (10), 1617-1630 (2000)  |
| MEDLINE   | 20499374   |
| PUBMED    | 11042159   |
| REFERENCE | 3  |
| AUTHORS   | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  |
| TITLE     | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer   |
| JOURNAL   | Genome Res. 10 (11), 1757-1771 (2000)  |
| MEDLINE   | 20530913   |
| PUBMED    | 11076861   |
| REFERENCE | 4  |
| AUTHORS   | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.   |
| TITLE     | Functional annotation of a full-length mouse cDNA collection   |
| JOURNAL   | Nature 409, 685-690 (2001)   |
| REFERENCE | 5  |
| AUTHORS   | The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.   |
| TITLE     | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs   |
| JOURNAL   | Nature 420, 563-573 (2002)   |
| REFERENCE | 6  |
| AUTHORS   | Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y. |
| TITLE     | Direct Submission  |
| JOURNAL   | Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  |
| COMMENT   | On Dec 10, 2002 this sequence version replaced gi:12851571. Please visit our web site (http://genome.gsc.riken.jp/) for further details.   |
| FEATURES  | cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGACGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGAGTATCTAAATTAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR. Location/Qualifiers  |

## FEATURES



```

QY 744 ATAAATTTTATCATATAATAAATAAATTAACCCCATACAGATCTGAAGCCGTGAAATAATTTTG 803
Db 653 ATAAATTTTATCATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 712
QY 804 TTTGTGAAGTCTGATCTAGTCAATATATTTCTAAATGAACGTTGATGAACGCA 863
Db 713 TTTGTGAAGTCTGATCTAGTCAATATATTTCTAAATGAACGTTGATGAACGCA 772
QY 864 CTGAAAAACACAGATATCAAAAGTTCTGTCACTTTGGAAGTGAACGTTGATGAACAT 923
Db 773 TTGAAAAACACAGATATCAAAAGTTCTGTCACTTTGGAAGTGAACGTTGATGAACAT 832
QY 924 CACAGTACTTTGGTCTACCCGCACTACAGAGTCCCGAGGTCAATTTGGCTTTAGT 983
Db 833 CATAGTACTTTGGTCTACCAAGGCACTACAGGCTCCAGAGGTCAATTTGGCTTAGT 892
QY 984 TGGTCTCAGCCTTGGATGTTTGGAGCATAGTTCATCTTATGAATATACCTTGGT 1043
Db 893 TGGTCTCAGCCTTGGATGTTTGGAGCATAGTTCATCTTATGAATATACCTTGGG 952
QY 1044 TTCACAGTCTTTCAGACTCATGATAGTAAGAGACCTGGCAATGATGAACGAATATTA 1103
Db 953 TTCACAGTCTTTCAGACCCACGATAGTAAGAGACCTGGCAATGATGAACGAATCTTA 1012
QY 1104 GGACCCATACCAACAACATGATTCAGAAAAACAAGAAACCGAAGTATTTTCCACATAAC 1163
Db 1013 GGACCCATCCCAAGCATATGATCCAGAAACAAGAAACCGAAGTATTTTCCACATAAC 1072
QY 1164 CAGCTAGATTTGGATGAACAGATCTCTGCTGGTAGATATGTTAGAGAGCGTGCNAACCG 1223
Db 1073 CAGCTAGATTTGGATGAACAGATCTCTGCTGGTAGATATGTTAGAGAGCGTGCNAACCG 1132
QY 1224 TTGAAGGAATTTATCTTTGTCATGATGAAGAAACATGAGAACTGTTTGACCTGGTTCGA 1283
Db 1133 TTGAAGGAATTTATCTTTGTCATGATGAAGAAACATGAGAACTGTTTGACCTGGTTCGA 1192
QY 1284 AGAATGTTAGATATGATTCACAACTCAAGAAATTAACCTTTGGATGAAGCAATTCAGCATCCT 1343
Db 1193 AGAATGTTAGATATGATTCACAACTCAAGAAATTAACCTTTGGATGAAGCAATTCAGCATCCT 1252
QY 1344 TTTCTTTGACTTATTAAGAAAGAAATGAATGGATCAGTGGTCTTACTATATCTTCTC 1403
Db 1253 TTTCTTTGACTTATTAAGAAAGAAATGAATGGATCAGTGGTCTTACTATATCTTCTC 1310
QY 1404 TAGAAGAGATTACTT-AAAGACTGTCTCAGTCAACTAAACATCTTAATATTTTGTAAACA 1462
Db 1311 TAGGAGCAGTTACTTCCAGACTGTCTCAGTCAACTAAACCTTCTAATATTTTGTAAACA 1370
QY 1463 TTAATATTTTGTACAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATTAATTA 1522
Db 1371 TTAATATTTTGTACAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATTAATTA 1430
QY 1523 ACTTGTTAAGCAAGTATCGTCTTGTATATGCAATTAAGAAAAATTAATTAATTTTCTT 1582
Db 1431 CCTTGTTAAGTATGGTGTGTAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1490
QY 1583 T-TGAAATTAACATTTTAAATACCTTTGAAATATCTTTGTGTCCAGTGAATTAATCTG 1640
Db 1491 TGATGTTAATTTGACACTTTAAGGGCTTTTGAATGCCCCCTTTGTGTCCAGTGAATTAAC 1550
QY 1641 ATTGATCTTGGCTTTTGTACATGGAGGTCACTCTGAAAGTGAATTTTTTTTTT-----GAG 1693
Db 1551 ATTGTTCCCATCTTTTGTACATGAAGGTGTGACTCTGAAAGTGAATTTTTTTTTTCTCTCAAG 1610
QY 1694 TAAAGGAAATCTTGAC 1710
Db 1611 TAAAGGAAATCTTGAC 1627

```

```

RESULT 6
AY408254
LOCUS

```

```

1421 bp DNA linear GSS 15-DEC-2003

```

```

DEFINITION Mus musculus CLK4 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY408254
VERSION AY408254.1 GI:39764225
KEYWORDS GSS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1421)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1421)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1421
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1421
/gene="CLK4"
/locus_tag="HCM3162"
ORIGIN
Query Match 41.6%; Score 978.2; DB 9; Length 1421;
Best Local Similarity 84.1%; Pred. No. 2.2e-191;
Matches 1081; Conservative 0; Mismatches 180; Indels 25; Gaps 1;
QY 85 GTCATTTATTAGAACGAGTCTTGAATGACGAGATATTCGGACCGGAGATACGTTG 144
Db 161 GTCATTTATTAGAACGAGATGCTTGAATGACGAGATATTCGGACCGGAGATACATTTG 220
QY 145 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATATATCACAGACATTTG 204
Db 221 ATGAATACAGAAATGACTACTGTGAAGGATATGTTCTCAAGACATTTACCATGAGACGTTG 280
QY 205 AAAGCGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCTCTA 264
Db 281 AAAGCACTTACCGATCCATTCGAGTAAATCTTCAGTCCGACGAGGAGAGAGCAGCCTTA 340
QY 265 AAAGGAGCGCAATAGACATCTGTTCAAGTCAATGATGACGTTCCGAGAGCCACCGAAGA 324
Db 341 AGAAGAACGTAATAGACCCCTGTGCAAGTCAATGATGACGTTCCGAGAGCCACCGAAGA 400
QY 325 AAAGTCCAGGATATAGAGGATGATGAGGAGGTCACTGATCTGTCAAGTGGAGAGC 384
Db 401 AAAGTCCAGGATATAGAGGATGATGAGGAGGTCACTGATCTGTCAAGTGGAGAGC 460
QY 385 TTCTAAGACGAGATATGAATTCGTGGACACTTTTGGGTGAAGAGCGCTTTGCGAAAGTTG 444
Db 461 TTCTAAGACGAGATATGAATTCGTGGAGTAAATTCGTGGAGTAAATTCGTGGAGTAAATTC 520
QY 445 TAGAGTGCATTTGATCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGG 504
Db 521 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 580
QY 505 TAGGCCGTTACCGTGAAGACGCTGTTTCAGAAATCCAAAGTATTAGAGCACTTAATAGTA 564
Db 581 TAGGCCGTTACCGGAGGAGCAGCTGTTTCTGAATCCAAAGTATTAGAGCACTTGAACAGCA 640

```

565 CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGAAATGTTTGTATCATCATGTC 624  
641 CTGACCCCAACAGTGTCTTCGATGTGTCAGATGCTAGAAATGTTTGTATCATCATGTC 700  
625 ATGTTTGTATGTTGTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 684  
701 ATGTTTGTATGTTGTTGAGCTGTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 760  
685 GCTTCTGCCATTTCAATTTGACACCATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 744  
761 GTTTTCTGCCATTTCAATTTGATCAGATCAGGCAATGGC----- 800  
745 TAAATTTTTCATCATATAATTAATTAACCCATACAGATCTGAAGCTCGAAATATTTTGT 804  
801 -----TTTTTACATCATATAATTAATTAACACGAGCTTAACCTGAAATATTTTAT 855  
805 TTGGAAGTCTGACTATGATGATCAATAATTAATTAACAAATGAAGCTGATGAACGACAC 864  
856 TTGGAAGTCTGACTATGATGATCAATAATTAATTAACAAATGAAGCTGATGAACGACAC 915  
865 TGAATAACACAGATATCAAGTGTGTTGACTTTGGAAGTGAACGATGATGATGAACATC 924  
916 TGAATAACACAGATATCAAGTGTGTTGACTTTGGAAGTGAACGATGATGATGAACATC 975  
925 ACAGTACTTTTGTGTTCTACCCGGCACTACAGAGTCTCCGAGGTCAATTTTGGCTTTAGGTT 984  
976 ATAGTACTTTTGTGTTCTACACAGGCACTACAGGCTCCAGAGGTCAATTTTGGCTTAGGTT 1035  
985 GGTCTCAGCTTTGATGTTTGGAGCATAGGCTGCTTCTTTTGAATATTAACCTTTGTTT 1044  
1036 GGTCTCAGCTTTGATGTTTGGAGCATAGGCTGCTTCTTTTGAATATTAACCTTTGTTT 1095  
1045 TCACAGTCTTTGAGTCTATGATGATGAAGACCTGGCCTGCAATGATGGAAGATATTAG 1104  
1096 TCACAGTCTTTGAGTCTATGATGATGAAGACCTGGCCTGCAATGATGGAAGATATTAG 1155  
1105 GACCCATACCAACACATGATTTTCAAGAAACGCAAGTATTTTCCACCATATACC 1164  
1156 GACCCATCCAGCATATGATTCAGAGACAGGAAACGCAAGTATTTTCCACCATATACC 1215  
1165 AGTAGATTGGATGAACACAGTCTTCTGCTGTAGATATGTTAGGAGACGCTGCAACCGT 1224  
1216 AGTAGATTGGAGCAGCATAGTTTCTGCTGTAGATATGTTAGGAGACGCTGCAACCGT 1275  
1225 TGAAGGATTTATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1284  
1276 TAAAGGAATTTATGCTGTGTCATGACGAAGAGCATGGAAGCTGTTGACCTGGTTCGAA 1335  
1285 GAATGTTTGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGACGATCCTT 1344  
1336 GAATGTTTGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGACGATCCTT 1395  
1345 TCTTTGACTTATTAAGAAATGA 1370  
1396 TCTTTGACTTATTAAGAAATGA 1421

RESULT 7  
AY408253  
LOCUS  
DEFINITION Pan troglodytes CLK4 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
AY408253  
VERSION AY408253.1 GI:39764224  
KEYWORDS GSS.  
ORGANISM Pan troglodytes (chimpanzee)  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
1 (bases 1 to 1446)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
Science 302 (5652), 1960-1963 (2003)  
14671302  
2 (bases 1 to 1446)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source location/Qualifiers  
1..1446  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>1446  
/gene="CLK4"  
/locus\_tag="HCM3162"  
ORIGIN  
Query Match 39,8%; Score 937.6; DB 9; Length 1446;  
Best Local Similarity 73.1%; Pred. No. 5.3e-183;  
Matches 940; Conservative 0; Mismatches 346; Indels 0; Gaps 0;  
Qy 85 GTCATTATTTAGAACGAGTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 144  
Db 161 GTCATTATTTAGAACGAGTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 220  
Qy 145 ACGAATACAGGAATGACTACTGTCAGAGATATGTTCTAGACATTTATCAGAGACATTTG 204  
Db 221 ATGAATACAGGAATGACTACTGTCAGAGATATGTTCTAGACATTTATCAGAGACATTTG 280  
Qy 205 AAACGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGAGAGAGACGATCCTA 264  
Db 281 AAACGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGAGAGAGACGATCCTA 340  
Qy 265 AAAGGAAGCGCAATAGACATCTGTTCAAGTCAATGATGATGATGATGATGATGATGATG 324  
Db 341 AAAGGAAGCGCAATAGACATCTGTTCAAGTCAATGATGATGATGATGATGATGATGATG 400  
Qy 325 AAAGTCCAGAGTATAGAGATGATGAGAGGTCACCTGATCTGTCAAAGTGGAGACG 384  
Db 401 AAAGTCCAGAGTATAGAGATGATGAGAGGTCACCTGATCTGTCAAAGTGGAGACG 460  
Qy 385 TTCTAAGAGCAAGATATGAAATCGTGACACTTTTGGGTGAAGAGCCCTTTGGCAAGTTG 444  
Db 461 NNN 520  
Qy 445 TAGAGTGCATTGATCATGGCATGATGATGATGATGATGATGATGATGATGATGATGATG 504  
Db 521 NNN 580  
Qy 505 TAGGCGGTACCGTGAAGCAGCTCGTTTCAGAAATCCAGATTTATAGAGCACTTTAAATAGTA 564  
Db 581 TAGGCGGTACCGTGAAGCAGCTCGTTTCAGAAATCCAGATTTATAGAGCACTTTAAATAGTA 640  
Qy 565 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGAAATGTTTGTATCATCATGTC 624  
Db 641 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGAAATGTTTGTATCATCATGTC 700  
Qy 625 ATGTTTGTATGTTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 684  
Db 701 ATGTTTGTATGTTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 760  
Qy 685 GCTTCTGCCATTTCAATTTGACACCATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 744  
Db 761 GCTTCTGCCATTTCAATTTGACACCATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 820

QY 745 TAAATTTTTCATCATATAATAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 804  
 Db 821 NNNNNNTTTTACATCATATAATAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 880  
 QY 805 TTGTGAAGTCTGACTATGTAGTCAATATAATTTAAATGAAGCTGATGAACGACAC 864  
 Db 881 NTNNNANGTCTGACTATGTNGTCAAAATAATAATTTCAAAATGAAGCTGATGAACGNNNN 940  
 QY 865 TGAATAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACGATGATGATGAACATC 924  
 Db 941 TGAATAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACGATGATGATGAACATC 1000  
 QY 925 ACAGTACTTTGTGTCTACCCGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGGTT 984  
 Db 1001 ACAGTACTTTGTGTCTACCCGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGGTT 1060  
 QY 985 GGTCTCAGCCTTGTGATGTTGGAGCATAGTTGCAATCTTATTTGAATATTTACCTTGGTT 1044  
 Db 1061 GGTNNCNGCCTTGTGATGTTGGAGCATAGTTGCAATCTTATTTGAATATTTACCTTGGTT 1120  
 QY 1045 TCACAGTCTTTTCAGACTCATGATAGTAAAGACCTGCAATGATGAACGAATATTAG 1104  
 Db 1121 TCACAGTCTTTTCAGACTCATGATAGTAAAGACCTGCAATGATGAACGAATATTAG 1180  
 QY 1105 GACCCATACCAACACATGATTTTCAGAAAAACAGAAAAACGCAAGTATTTTCAACATAACC 1164  
 Db 1181 GACCCATACCAACACATGATTTTCAGAAAAACAGAAAAACGCAAGTATTTTCAACATAACC 1240  
 QY 1165 AGCTAGATGGGATGAACACAGTCTTCGTGTGTAGATATGTTAGGAGAGCTGCAACCGT 1224  
 Db 1241 NNN 1300  
 QY 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTGACCTGGTTCGAA 1284  
 Db 1301 TGAAGNN 1360  
 QY 1285 GAAATGTAAGATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCT 1344  
 Db 1361 NNT 1420  
 QY 1345 TCTTTGACTTATTAATAAAGAAATGA 1370  
 Db 1421 TCTTTGACTTATTAATAAAGAAATGA 1446

RESULT 8  
 CR591789  
 LOCUS  
 DEFINITION full-length cDNA clone CL0BA009ZE09 of Placenta of Homo sapiens (human).  
 ACCESSION CR591789  
 VERSION CR591789.1 GI:50472596  
 KEYWORDS HTC; CNSLT CDNA.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1737)  
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalisation  
 JOURNAL Unpublished  
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue  
 REFERENCE 2 (bases 1 to 1737)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)  
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dn) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES  
 source  
 1..1737  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CL0BA009ZE09"  
 /tissue\_type="Placenta"  
 /plasmid="pCMVSPORT\_6"

## ORIGIN

Query Match 34.6%; Score 814.8; DB 3; Length 1737;  
 Best Local Similarity 73.4%; Pred. No. 1.2e-157;  
 Matches 1102; Conservative 0; Mismatches 382; Indels 18; Gaps 4;

QY 84 AGTCATTTATTTAGAAAGCAAGGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTT 143  
 Db 238 AGCCATTTATTTGGAAAGCAGGTCTATAAATGAGAAGATTATCATAGTCGACGCTACATT 297  
 QY 144 GACGAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTATACAGAGACATT 203  
 Db 298 GATGAGTACAGAAATGACTACCTCAAGGATGTGAACCTGGACATCGCCAAAGAGACCAT 357  
 QY 204 GAAAGCGGTATCGAATCCACTGCAGTAAATCTTTCAGTCCGACGAGGAGAGCAGTCCT 263  
 Db 358 GAAAGCGGTATCAGAACCATAGTAGCAAGTCTTCTGGTAGAAGTGAAGAAGTAGTTAT 417  
 QY 264 AAAAGGAAGCGCAATAGACAC-----TGTTCAAGTCAATAGTCACGTTCCGAAGAGCCAC 317  
 Db 418 AAAAGCAACACAGGATTTCAACAGACTCTTCACATCGTCCTTCACATGGGAGAGTCAC 477  
 QY 318 CGAAGGAAGATCCAGGATATAGAGGATGATGAGGAGGCTCACCTGATCTGTCAAGT 377  
 Db 478 CGAAGGAAGAAACACAGGATGTAGAGGATGATGAGGAGGCTCACCTGATCTGTCAAGT 537  
 QY 378 GGAGACGTTCTTAAGAGCAAGATATGAATCCGTGGACACTTTTGGGTGAAGAGCCCTTTGGC 437  
 Db 538 GGAGACGTTACTAAGTGCAGATATGAATTTGTGTACTTTAGGTGAAGAGGCTTTTGA 597  
 QY 438 AAAGTTGTAGAGTGCATGATGCGCATGATGCGCATGATGCGCATGATGCGCATGATGCGTA 497  
 Db 598 AAAGTTGTGAGTGCATGATGCGCATGATGCGCATGATGCGCATGATGCGCATGATGCGTA 657  
 QY 498 AAAATGTAGCCGTTACCGTGAAGCAGCTGTTTCAGAAATCCAGATTTAGAGCAGCTTA 557  
 Db 658 AAAATGTGAGTATGATCTGTGAAGTCTGCTCGCTCAGAAATACAAAGTTCTGGAACTCTG 717  
 QY 558 AATAGTACTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGATGTTGTTGATCAT 617  
 Db 718 AATACACAGACCCCAACAGACTCTTTCGCTGTGTCCAGATGTTGGAATGTTGAGCAT 777  
 QY 618 CATGCTCATGTTTGTATTTGTTGAACTACTCGGACTTACTTACGATTTCAATTTAA 677  
 Db 778 CATGCTCATGTTTGTATTTGTTGAACTACTCGGACTTACTTACGATTTCAATTTAA 837  
 QY 678 GAAACAGCTTTCTGCCATTTCAATTCACACATCAGGAGATGCGGATGCGGATCAGATCTGC 737  
 Db 838 GAAATGTTTCTTACCATTTCGACTGGATCATATCAAGAAAGTGGCATATCAGATATGCG 897  
 QY 738 CAGTCAATAAATTTTATACATATAAATAAATTAACCCATACAGATCTGAAGCCTGAAAT 797  
 Db 898 AAGTCTGTGAATTTTTCACAGTAATAAGTTGACTCACACAGACTTAAGCCTGAAAC 957  
 QY 798 ATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTAAATGAACGATGAA 857  
 Db 958 ATCTTATTTGTGAGTCTGACTACACAGAGCGGTATAATCCCAAAATAAAGCGATGAA 1017  
 QY 858 CGCACACTGAAACACAGATATCAAGTGTGTGACTTTGGAAGTGCACCTATGATCAT 917  
 Db 1018 CGCACCTTAATAATCCAGATATTAAGTGTGTAGACTTTGGTAGTGCACATATGATGAC 1077

|            |  |   |                             |
|------------|--|---|-----------------------------|
| Qy         | 918  | GAACATCAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCGAGGTCAATTTGGCT       | 977                         |
| Db         | 1078   | GAACATCAGTACTTTGGTGTCTACCAAGACATTTAGAGCACCTGAGTTATTTTGGC        | 1137                        |
| Qy         | 978  | TTAGTTGGTCTCAGCCTTGTGATGTTTGGAGCATAGTGGTCTTATTTGAATATAC         | 1037                        |
| Db         | 1138   | CTAGGGTGGTCCCAACCATGTGTCTGAGGACATAGATGCAATTTTATTTGAATATAT       | 1197                        |
| Qy         | 1038   | CTTGGTTTCAAGTCTTTTCAAGTCAATGATAGTAAAGAGCACCTGGCAATGATGAACGA     | 1097                        |
| Db         | 1198   | CTTGGTTTACCGTATTTTCCAAACACATGATAGTAAAGAGCATTTAGCAATGATGAACAG    | 1257                        |
| Qy         | 1098   | ATATTAGGACCCATACCAACACATGATTTAGAAACCAAGAAACGCAAGTATTTTCCAC      | 1157                        |
| Db         | 1258   | ATTTCTGGACCTCTACCAAAACATATGATACAGAAACCCAGAAACGTAATATTTTCCAC     | 1317                        |
| Qy         | 1158   | CATTAACAGCTAGATTGGATGAACACAGTTCTGCTGTGTAGATATTTAGAGAGCGTGC      | 1217                        |
| Db         | 1318   | CACGATCATTTAGATGGGATGAACACAGTTCTGCGGACATATGTTTCAAGACGCTGT       | 1377                        |
| Qy         | 1218   | AAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAACTGTTTGACCTG    | 1277                        |
| Db         | 1378   | AAACCTCTGAAGGAATTTATGCTTTTCTCAAGATGTTGAACATGAGCGTCTCTTTGACCTC   | 1437                        |
| Qy         | 1278   | GTTTGAAGAAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCAG | 1337                        |
| Db         | 1438   | ATTTCAGAAATGTTGGAGTATGATCCAGCAAAAGAAATTTACTCTCAGAGAAACCTTAAAG   | 1497                        |
| Qy         | 1338   | CATCTTTCTTTGACTTATTTAAAGAAATGAAATGGAAATCAGTGTCTTACTATATA        | 1397                        |
| Db         | 1498   | CATCTTTCTTTGACCTTCTGAAGAAAA-----GTATATAGATCTGTAATTTGGACA        | 1548                        |
| Qy         | 1398   | CTTCTCTAGAAGAGATTACTTAAAGACTGTGTCAGTC--AACTAAACATTTCTAAATATTTT  | 1455                        |
| Db         | 1549   | GCTCTCTCGAAGAGA-TCTTACAGACTGTATCAGTCTAATTTTAAATTTTAAAGTATTT     | 1607                        |
| Qy         | 1456   | GTAACATTTAAATTTTGTACAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGC        | 1515                        |
| Db         | 1608   | TGTACAGCTTTGTAAATTTCTTAAACATTTTATTTATTTGCTTATTTTGTGGTAA         | 1667                        |
| Qy         | 1516   | ATAATTAATTTGTTAAAGCAAGTATGCTTGTATGATGATGATGATGATGATGATGATG      | 1575                        |
| Db         | 1668   | TTGGTTTCAATTAAGTATAGCTAAGGTAAGTAAAGCAATCTTTTTCAGTAATTTGTAAGTAT  | 1727                        |
| Qy         | 1576   | TT  | 1577                        |
| Db         | 1728   | TT  | 1729                        |
| RESULT 9   |  |   |                             |
| LOCUS      | CR612115   | 1746 bp   | mRNA linear HTC 21-JUL-2004 |
| DEFINITION | full-length cDNA clone CS0DF036YH24 of Fetal brain of Homo sapiens (human).  |   |                             |
| ACCESSION  | CR612115   |   |                             |
| VERSION    | CR612115.1 GI:50492922   |   |                             |
| KEYWORDS   | HTC; CNSLT cDNA.   |   |                             |
| SOURCE     | Homo sapiens (human)   |   |                             |
| ORGANISM   | Homo sapiens   |   |                             |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.           |   |                             |
| AUTHORS    | 1 (bases 1 to 1746)  |   |                             |
| TITLE      | Li.W.B., Gruber,C., Jesse,J. and Polayes,D.  |   |                             |
| JOURNAL    | Full-length cDNA libraries and normalization   |   |                             |
| REMARK     | Unpublished  |   |                             |
|            | Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue    |   |                             |
| REFERENCE  | 2 (bases 1 to 1746)  |   |                             |
| AUTHORS    | Genoscope.   |   |                             |
| TITLE      | Direct Submission  |   |                             |
| JOURNAL    | Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr |   |                             |

|          |   |   |      |
|----------|---|---|------|
| COMMENT  | - Web : www.genoscope.cns.fr)<br>1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. |   |      |
| FEATURES | Location/Qualifiers   |   |      |
| source   | 1..1746   |   |      |
|          | /organism="Homo sapiens"  |   |      |
|          | /mol_type="mRNA"  |   |      |
|          | /db_xref="taxon:9606"   |   |      |
|          | /clone="CS0DF036YH24"   |   |      |
|          | /tissue_type="Fetal brain"  |   |      |
|          | /plasmid="pCMVSPORT_6"  |   |      |
| ORIGIN   |   |   |      |
|          | Query Match   | 34.6%; Score 814.8; DB 3; Length 1746;                        |      |
|          | Best Local Similarity   | 73.4%; Pred. No. 1.2e-157; Indels 18; Gaps 4;                 |      |
|          | Matches 1102; Conservative  | 0; Mismatches 382;  |      |
| Qy       | 84  | AGTCATTATTTAGAAAGCAAGGTCTTGAATCAGCGAGATTATCGGGACCGGAGATACGTT  | 143  |
| Db       | 237   | AGCCATTATTTGGAAGCAGGTCTATAATGAGAAGATTATCATAGTCGACGCTACATT     | 296  |
| Qy       | 144   | GACGAATACAGAAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGACATT   | 203  |
| Db       | 297   | GATGATACAGAAATGACTACACTCAAGGATGTGAACCTGGACATCGCCAAAGAGACCAT   | 356  |
| Qy       | 204   | GAAGCGGGTATCGAATCCACTGCGAGTAAATTTTCAGTCCGACGAGAGAGAGCTCT      | 263  |
| Db       | 357   | GAAGCGGGTATCAGAACCATAGTAGCAAGTCTTCTGTAGAGTGGAGAAGTAGTTAT      | 416  |
| Qy       | 264   | AAAAGGAAGCGCAATAGACAC-----TGTTCAAGTCATCAGTCACGTTTCGAAGAGCCAC  | 317  |
| Db       | 417   | AAAAGCAACACAGATTTCACACAGTACTTTCACATGCTGCTTCACATGGGAAGAGTCAC   | 476  |
| Qy       | 318   | CGAAGGAAGATCCAGAGTATAGAGTATGATGAGGAGGTGACCTGATCTGTCAAAAGT     | 377  |
| Db       | 477   | CGAAGGAAGAAAGCAGAGGTGTAGAGGATGATGAGGAGGTGACCTGATCTGTCAAGT     | 536  |
| Qy       | 378   | GGAGACGTTCTTAAGAGCAAGATATGAATCTGTGACACCTTTGGGTGAAGAGCGCTTTGGC | 437  |
| Db       | 537   | GGAGACGTTCTTAAGAGCAAGATATGAATCTGTGACACCTTTGGGTGAAGAGCGCTTTGGC | 596  |
| Qy       | 438   | AAAGTTGTAGAGTGCATTGATCATGGCATGGATGGCATGGCATGGCATGGATGGATGGAT  | 497  |
| Db       | 597   | AAAGTTGTGGAGTGCATCGATCATAAAGCGGAGGTAGACATGATGACATGATGACATGAT  | 656  |
| Qy       | 498   | AAAAATGTAGCCGTTACCGTGAAGCAGCTGTTTCAGAAATCCAAAGTATTAGAGCACCTTA | 557  |
| Db       | 657   | AAAAATGTGGATAGATATCTGTGAAGCTGCTCGCTCAGAAATACAAAGTTCTGGAACTCTG | 716  |
| Qy       | 558   | AATAGTACTGATCCCAATAGTGTCTTCGATGTGTCCAGATGTCAGATGTTGTTGATCAT   | 617  |
| Db       | 717   | AATACACAGACCCCAACAGTACTTTTCGCTGTGTCCAGATGTTGGAATGTTGAGCAT     | 776  |
| Qy       | 618   | CATGTCATGTTTGTGTTTGTGTTTGAACACTCTGGGACTTAGTACTTACGATTTCAATTA  | 677  |
| Db       | 777   | CATGTCATGTTTGTGTTTGTGTTTGAACACTCTGGGACTTAGTACTTACGATTTCAATTA  | 836  |
| Qy       | 678   | GAAGACAGCTTTTCGCCATTTTCAAAATTCACCATCAGGACATGCGCTATCAGATCTGC   | 737  |
| Db       | 837   | GAAGACAGCTTTTCGCCATTTTCAAAATTCACCATCAGGACATGCGCTATCAGATCTGC   | 896  |
| Qy       | 738   | CAGTCAATAAAATTTTTCATCATATAATAAATTAACCCATACAGATCTCAAGCGCTGAAAT | 797  |
| Db       | 897   | AAGTCTGCAATTTTTCGACAGTAATAAGTTGACTCACACAGACTTAAAGCGCTGAAAC    | 956  |
| Qy       | 798   | ATTTTGTGTTGAAAGTCTGACTATGTTAGTCAAAATATAAATTTTAAATGAAGCGTGA    | 857  |
| Db       | 957   | ATCTTATTTGTGAGTCTGACTACACAGAGGCGTATATATCCCAAAATAAAGCGTGA      | 1016 |
| Qy       | 858   | CGCACACTGAAACACAGATATCAAAGTTGTTGAGTCTTGGAGTGCACCGTATGATGAT    | 917  |

```
Db 1017 CGCACCTTAATAAATCCAGATATTAAAGTTGTAGACTTTGGTAGTGCACATATGATGAC 1076
Qy 918 GAAACATACAGTACTTTGTGTCTACCGGCACTACAGAGCTCCGAGGTCAATTTGGCT 977
Db 1077 GAACATACAGTACTTTGTGTCTACCGGCACTACAGAGCTCCGAGGTCAATTTGGCT 1136
Qy 978 TTAGTGTGCTTCAGCTTTGTGATGTTTGGAGCATAGGTTCATTTATTGAATATTAC 1037
Db 1137 CTAGGGTGTGCTCCCAACCATGTCTGTGGAGCATAGGATGATCTTTATTGAATATTAT 1196
Qy 1038 CTTGGTTTACAGTCTTTTCAGACTCATGATAGTAAGAGCACCCTGGCAATGATGAAAGA 1097
Db 1197 CTTGGGTTTACCGTATTTTCCACACACAGCATAGTAAGAGGCAATTTAGCAATGATGAAAG 1256
Qy 1098 ATATTAGAGCACCATACCAACACATGATTTCCAGAAAACAGAAAACGCAAGTATTTCAC 1157
Db 1257 ATCTTGGACCTCTACCAAAACATATGATACAGAAAACAGGAAACGTAATATTTCAC 1316
Qy 1158 CATAACCACTAGATGTTGGATGAACACACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGC 1217
Db 1317 CACGATCGATTAGACTGGGATGAACACAGTCTCTGCGGCAGATATGTTCAAGACGCTGT 1376
Qy 1218 AAACCGTTGAAGGAATTTATGCTTTGTCTCATGATGAAGAAACATGAGAAAACGTTTGAACCTG 1277
Db 1377 AAACCTCTGAAGGAATTTATGCTTTCTCAAGATGTTGAACATGAGCGTCTCTTTGACCTC 1436
Qy 1278 GTTCCAGAAATGTTAGATATGATCCAACTCAAGAAATTTACCTGGATGAAGCAATGCGAG 1337
Db 1437 ATTCAAGAAATGTTGGAGATGATCCAGCCAAAGAAATTTACTCTCAGAGAACGCTTTAAAG 1496
Qy 1338 CATCTCTTTTGTGACTTATTAAGAAATGAAATGGAATCAGTGTCTTACTATATA 1397
Db 1497 CATCTCTTTTGTGACTTCTGAGAAA-----GTATATAGATCTGTAATTGGACA 1547
Qy 1398 CTTCTCTAGAGAGATTTACTTAAGACTGTGTCAGTC--AACTAAACATTTCTAATTTT 1455
Db 1548 GCTCTCTCGAAGAGA-TCTTACAGACTGTATCAGTCTAATTTTAAATTTTAAAGTTATTT 1606
Qy 1456 GTAACATTAATTTTGTGACAGTAAAGTAAATTTATGATGTTTGTGATCAATAGC 1515
Db 1607 TGTAAGCTTTGTAATTTCTTAAACATTTTATATTTGTCATGTTTATTTTGTGGTAA 1666
Qy 1516 ATAATTAATTTGTTAAGCAAGTATGCTCTGATATGATGATGATGATGATGATGATGAT 1575
Db 1667 TTGGTTTCAATTAAGTACATAGCTAAGGTAATGAACATCTTTTCAGTAATTTGTAAGTAT 1726
Qy 1576 TT 1577
Db 1727 TT 1728

RESULT 10
CR617472
LOCUS
DEFINITION
full-length cDNA clone CS0D1036XH19 of Placenta linear HTC 21-JUL-2004
of Homo sapiens (human).
ACCESSION
CR617472.1 GI:50498279
VERSION
HTC; CDSLT cDNA.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1754)
AUTHORS
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
REFERENCE
2 (bases 1 to 1754)
AUTHORS
Genoscope.
```

```
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1754
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1036XH19"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 34.6%; Score 814.8; DB 3; Length 1754;
Best Local Similarity 73.4%; Pred. No. 1.2e-157;
Matches 1102; Conservative 0; Mismatches 382; Indels 18; Gaps 4;
Qy 84 AGTCATTATTTAGAGCAAGGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTT 143
Db 238 AGCCATTATTTGGAAAGCAGGTCTATAAATGAGAAAGATTATCATAGTCGACGCTACATT 297
Qy 144 GACCAATACAGGAATGACTACTGTGAGGATATGTTCTAGACATTATCACAGAGACATT 203
Db 298 GATGAGTACAGAAATGACTACACTCAAGGATGTGAACCTGGACATCGCCAAAGAGACCAT 357
Qy 204 GAAAGCGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGAGCAGGAGAGAGCTCT 263
Db 358 GAAAGCGGTATCAGAACCATAGTAGCAAGTCTTCTGTAGAGTGGAGAGAGTAGTTAT 417
Qy 264 AAAAGGAAGCCAAATAGACAC-----TGTTCAAAGTCAATAGTCACGTTTCGAAGAGCCAC 317
Db 418 AAAAGCAACACAGAGTTTCCACACAGTACTTTCACATCGTCTGTTACATCGGAGAGAGTCAC 477
Qy 318 CGAAGGAAGAGATCCAGAGTATAGAGGATGATGAGGAGGCTCACCTGATCTGTCAAAGT 377
Db 478 CGAAGGAAGAGAACACAGAGTGTAGAGGATGATGAGGAGGCTCACCTGATCTGTCAAAGT 537
Qy 378 CGAGACGTTTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGGCTTTGGC 437
Db 538 CGAGACGTTTCTAAGTGCAGATATGAAATCTTGTATCTTTAGGTGAAGAGGCTTTTGA 597
Qy 438 AAAGTTGTAGAGTGCATGATCATGGGCATGATGGCATGATGATGATGATGATGATGATGAT 497
Db 598 AAAGTTGTAGAGTGCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 657
Qy 498 AAAAATGTAGCGCTTACCGTGAAGCAGCTGTTCCAGAAATCCAGTATTAGAGCATT 557
Db 658 AAAAATGTAGGATAGATATCTGTGAAGCTGCTCGCTCAGAAATACAAAGTTCTGGAACATCTG 717
Qy 558 AATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGATGCTGTTGATCAT 617
Db 718 AATACACAGAGCCCAACAGTACTTTCCGCTGTGTCAGATGTTGGATGTTGGAGCAT 777
Qy 618 CATGCTCATGTTTGTATTTGTTTGAATCTACTGGGACTTAGTACTTACGATTTCAATTA 677
Db 778 CATGCTCATATTTGATTTGTTTGAATCTACTGGGACTTAGTACTTACGATTTCAATTA 837
Qy 678 GAAACAGCTTTCTGCCATTTCAATTTGACACATCAGGAGAGTGGGTATCAGATCTGCG 737
Db 838 GAAATGCTTTTCTTACCATTTTCGACTGATCATATCAAGAAAGATGGCATATCAGATATGCG 897
Qy 738 CAGTCAATAAATTTTTCATCATATAAATAAATTAACCCATACAGATCTCAAGCCTGAAAT 797
Db 898 AAGTCTGTGAATTTTTCGACAGTAATAGTTGACTTACACAGACTTAAAGCCTGAAAC 957
Qy 798 ATTTTGTGTGAAGTCTGACTATGATGATCAAAATATAATTTTAAATGAAACGATGAA 857
```



|            |  |   |      |
|------------|--|---|------|
| Db         | 666  | GAAGCGCAAGTATTTCCACCATACACCGTAGATTGGACGAGCATAGTTCAGCTGGGA   | 725  |
| Qy         | 1198   | GATATGTTAGGAGAGCGCTGCAAAACCGTTGAAGGAATTTATGCTTGTGTCATGATGAAGAAC   | 1257 |
| Db         | 726  | GATATGTTAGGAGAGCGCTGCAAGCGTTAAGGAATTTATGCTGTCATGACGAGAGC  | 785  |
| Qy         | 1258   | ATGAGAACTGTTTGACCTGTTGCGAAGATGTTAGATATGATCCAACTCAAGAAATTA   | 1317 |
| Db         | 786  | ATGAGAAGCTGTTTGACCTGTTGTTT-AAAAATGTTGGAGTATGATCACCACCGAGAGATCA  | 844  |
| Qy         | 1318   | -CCTTGGATGAGCATTCAGCATCCTTCTTGTGACTTATTAAAGAAAGAAATGAATCGG  | 1376 |
| Db         | 845  | TCTTTGGATGAGCATTCAGCATCCTTCTTGTGACTTATTAAAGAAAGAAATGAATCGG  | 904  |
| Qy         | 1377   | AATCAGTGTCTTACTATATATCTCT-AGAAGAGATTAAGCATGTCAGTCAAA  | 1435 |
| Db         | 905  | AGCCAGGGGCTTCGGGAACCTCCCTAAGACAGGTACTCCAGATGGGTACGCCAA  | 964  |
| Qy         | 1436   | CTAAACATT 1444  |      |
| Db         | 965  | CTAAACCTT 973   |      |
| RESULT 12  |  |   |      |
| AK076199   |  |   |      |
| LOCUS      |  |   |      |
| DEFINITION | AK076199   | 1689 bp mRNA linear HTC 03-APR-2004   |      |
|            |  | Mus musculus 18 days pregnant adult female placenta and extra   |      |
|            |  | embryonic tissue cDNA, RIKEN full-length enriched library, clone:3830430K12 product:PROTEIN KINASE CLK1 (EC 2.7.1.1-) (CLK) |      |
| ACCESSION  | AK076199   |   |      |
| VERSION    | AK076199.1   | GI:26096662   |      |
| KEYWORDS   | HTC; CAP trapper.  |   |      |
| SOURCE     | Mus musculus (house mouse)   |   |      |
| ORGANISM   | Mus musculus   |   |      |
| REFERENCE  |  |   |      |
| AUTHORS    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Carninci, P. and Hayashizaki, Y.   |   |      |
| TITLE      | High-efficiency full-length cDNA cloning   |   |      |
| JOURNAL    | Mech. Enzymol. 303, 19-44 (1999)   |   |      |
| MEDLINE    | 99279253   |   |      |
| PUBMED     | 10349636   |   |      |
| REFERENCE  |  |   |      |
| AUTHORS    | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.   |   |      |
| TITLE      | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes   |   |      |
| JOURNAL    | Genome Res. 10 (10), 1617-1630 (2000)  |   |      |
| MEDLINE    | 20499374   |   |      |
| PUBMED     | 11042159   |   |      |
| REFERENCE  |  |   |      |
| AUTHORS    | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsuoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. |   |      |
| TITLE      | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer   |   |      |
| JOURNAL    | Genome Res. 10 (11), 1757-1771 (2000)  |   |      |
| MEDLINE    | 20530913   |   |      |
| PUBMED     | 11076861   |   |      |
| REFERENCE  |  |   |      |
| AUTHORS    | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.   |   |      |
| TITLE      | Functional annotation of a full-length mouse cDNA collection   |   |      |
| JOURNAL    | Nature 409, 685-690 (2001)   |   |      |
| REFERENCE  |  |   |      |
| AUTHORS    | The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.   |   |      |

|                       |  |
|-----------------------|--|
| TITLE                 | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs   |
| JOURNAL               | Nature 420, 563-573 (2002)   |
| REFERENCE             | 6 (bases 1 to 1689)  |
| AUTHORS               | Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsumi, N., Okazaki, Y., Saigo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. |
| TITLE                 | Direct Submission  |
| JOURNAL               | Submitted (15-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]  |
| COMMENT               | cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/  |
| FEATURES              | Location/Qualifiers  |
| source                | 1. 1689  |
|                       | /organism="Mus musculus"   |
|                       | /mol_type="mRNA"   |
|                       | /strain="C57BL/6J"   |
|                       | /db_xref="FANTOM:DB:3830430K12"  |
|                       | /db_xref="taxon:10090"   |
|                       | /clone="3830430K12"  |
|                       | /sex="female"  |
|                       | /tissue_type="placenta and extra embryonic tissue"   |
|                       | /clone_lib="RIKEN full-length enriched mouse cDNA library"   |
|                       | /dev_stage="18 days pregnant adult"  |
|                       | 105. .1556   |
| misc_feature          | /note="PROTEIN KINASE CLK1 (EC 2.7.1.1-) (CLK) (PROTEIN KINASE STY) homolog [Mus musculus] (SWISSPROT P22518, evidence: FASTY, 95%ID, 100%length, match=1449) putative"  |
| polyA_signal          | 1670. .1675  |
|                       | /note="putative"   |
| polyA_site            | 1689   |
|                       | /note="putative"   |
| ORIGIN                |  |
| Query Match           | 31.2%; Score 734.2; DB 3; Length 1689;   |
| Best Local Similarity | 74.2%; Pred. No. 5.5e-141;   |
| Matches               | 957; Conservative 0; Mismatches 324; Indels 9; Gaps 2;   |
| Qy                    | 83 CAGTCATTATTAGAACGAGTCTTGAATGAGGAGATTATCGGACCGAGATACGT 142   |
| Db                    | 263 CAGCTATTATCTGGAAAGCAGATCCATAAATGAGAAAGCTTATCATAGTCGACGCTATGT 322   |
| Qy                    | 143 TGACGAATACAGGAATGACTACTCTGAAGGATATGTTCTTAGACATTATTCACAGACAT 202  |
| Db                    | 323 TGATGAATACAGGAATGACTATCATG---GGTACAGCCAGGATCCCTTGGAGAAC 379  |
| Qy                    | 203 TGAAGCGGGTATCGAATCCACTGCAGTAATCTTTCAGTCCGACGAGAGAGAGTCC 262  |
| Db                    | 380 TGAAGCAGATACCATGTCATAGTACGAGTCTCTGGTAGGAGTGAAGAGCAGTTA 439   |
| Qy                    | 263 TAAAGGAGGCGCAATAG-----ACACTGTTCAAGTCATCAGTCACCTTCGAGAGCCA 316  |



440 ACAAGTAAACACAGGAGTCGCCACACACTTCGCGAGCACCAATTCACACGGGAGAGTC 499  
 317 CCGAAGGAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTCTCAAG 376  
 500 CCGAAGGAATTAAGCGAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTCTCAAG 559  
 377 TGGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGGAGCTTTGG 436  
 560 TGGAGAGCTTCTAAGTCAAGATATGATATGTTGATACCTTTAGTGAAGGTCCTTCGG 619  
 437 CAAGTGTGTAGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 496  
 620 AGAAGTGTGGAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 679  
 497 AAAAAAGTGTAGCGCTTACCTGAGGAGGAGCTGTTCAAGAAATCAAGTATTAGAGCACT 556  
 680 TAAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739  
 557 AAATAGTACTGATCCCAATAGTCTTCCGATGTTCCGATGTTCCGATGTTCCGATGTT 616  
 740 GTATACAGCAGACCCCATAGTACTTCCGTTGGTGTCCAGATGTTGGATGTTGATGA 799  
 617 TCATGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 676  
 800 TCGAGATCACATTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 859  
 677 AGAAAACAGCTTCTGCGATTTCAAAATTTGACCATGACGAGGAGGATGCGTATCAGATCTG 736  
 860 GGAACACAGTTTCTGCGGTTCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 919  
 737 CCAGTCAATAAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCCTGAAAA 796  
 920 CAATGCTGTAACCTTTTGCATAGTAAATTAATTCACATACAGATCTGAAGCCTGAAAA 979  
 797 TATTTTGTGTAAGTCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 856  
 980 CATCTTATTTGTAAGTCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 1039  
 857 ACGACACTGAAAAACACAGATATCAAAATTTGTTGATCTTGGAGTGCACGATGATGA 916  
 1040 ACGTACTATAGTAATCCAGATATTAAGTGTGATGTTGATGTTGATGTTGATGTTGATGTT 1099  
 917 TGAACATCACAGTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 976  
 1100 TGAACACACAGCATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1159  
 977 TTTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1036  
 1160 CTTGGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1219  
 1037 CTTGGTTCAGTCTTTTCAAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1096  
 1220 TCTTGGATTTACAGTTTTCGATCTCATGATGATGATGATGATGATGATGATGATGATGATG 1279  
 1097 AATATTAGGACCCATACCAACATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1156  
 1280 GATTCTTGGACCACTACCAAGGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1339  
 1157 CCATAACAGCTAGATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1216  
 1340 TCATGATCGATTAGATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1399  
 1217 CAACCCGTTGAAGGAATTTATGCTTTTGTATGATGATGATGATGATGATGATGATGATGATG 1276  
 1400 TAACTCTGAAGGATTTATGCTATCTCAGATGATGATGATGATGATGATGATGATGATGATGAT 1459  
 1277 GGTTCGAAGATGTTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1336  
 1460 CATTTGGGAAAAATGTTGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1519  
 1337 GCATCCCTTTCTGATCTTTTAAAAAGAA 1366  
 1520 GCATCCCTTTCTTTTACCCTTTTAAAAAGCA 1549

RESULT 13

BM786882

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BM786882 722 bp mRNA linear EST 05-MAR-2002

K-EST0065716 S10SNUI Homo sapiens cDNA clone S10SNUI-17-D07 5',

mRNA sequence.

BM786882

BM786882.1 GI:19135114

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 722)

Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,

Oh, K.J., Cheong, J.B., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and

Kim, Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yusong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 17 row: D column: 07

High quality sequence stop: 722.

Location/Qualifiers

1..722

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S10SNUI-17-D07"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Lymphoblast-like"

/cell\_line="SNU-1"

/lab\_host="Top10F"

/clone\_lib="S10SNUI"

/notes="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;

Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then dephosphorylated

with tobacco acid pyrophosphatase (TAP). The dephosphorylated

intact mRNA was ligated with DNA-RNA linker including SfiI

site by treatment of T4 RNA ligase and the first strand

cDNA was synthesized with Superscript II using SfiI

oligo-dT primer. After first strand synthesis, RNA was

degraded by NaOH treatment and cDNA was amplified by PCR

reaction. The PCR products were digested with SfiI and

cloned into DrII- digested pME18S-FL3 vector. The

obtained cDNA vectors were used for transfection of

competent cells E. coli Top10F' by electroporation method.

The cDNA libraries constructed by this method are

full-length enriched cDNA library."

ORIGIN

Query Match

Best Local Similarity

Matches 722; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

682 ACAGCTTTCTGCCATTTCAAATTTGACCATCATCAGGAGATGGCGTATCAGATCTGCCAGT 741

1 ACAGCTTTCTGCCATTTCAAATTTGACCATCATCAGGAGATGGCGTATCAGATCTGCCAGT 60

742 CAATAAATTTTACATCATATAAATTAACCATCAGATCTGAAGCCTGAAAAATTTT 801

61 CAATAAATTTTACATCATATAAATTAACCATCAGATCTGAAGCCTGAAAAATTTT 120

802 TGTTTTGGAAGTCTGACTATGTAGTCAATATATATTTCTAAATGAAACGATGAAACGCA 861

121 TGTTTTGGAAGTCTGACTATGTAGTCAATATATATTTCTAAATGAAACGATGAAACGCA 180







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 07:43:58 ; Search time 393.184 Seconds  
(without alignments)  
9796.410 Million cell updates/sec

Title: US-10-801-671-1  
Perfect score: 2354  
Sequence: 1 gccagctgggtacttta.....taaaaaacaaaaaaac 2354

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description        |
|------------|--------|-------------|--------|----|--------------------|
| 1          | 2354   | 100.0       | 2354   | 3  | US-09-810-671-1    |
| 2          | 2354   | 100.0       | 2354   | 4  | US-10-109-854-1    |
| 3          | 2354   | 100.0       | 2354   | 4  | US-10-339-656-1    |
| 4          | 2052.6 | 87.2        | 2446   | 2  | US-09-016-000-9    |
| 5          | 1178.2 | 50.1        | 1456   | 4  | US-09-023-655-699  |
| 6          | 1097.4 | 46.6        | 1549   | 4  | US-09-905-999-26   |
| 7          | 1097   | 46.6        | 21234  | 3  | US-09-810-671-3    |
| 8          | 1097   | 46.6        | 21234  | 4  | US-10-109-854-3    |
| 9          | 1097   | 46.6        | 21234  | 4  | US-10-339-656-3    |
| 10         | 640.6  | 27.2        | 2254   | 4  | US-09-919-039-238  |
| 11         | 481.2  | 20.4        | 1538   | 4  | US-09-905-999-22   |
| 12         | 454.8  | 19.3        | 1787   | 4  | US-09-905-999-24   |
| 13         | 432.4  | 18.4        | 1762   | 4  | US-09-016-434-1439 |
| 14         | 430.8  | 18.3        | 1763   | 4  | US-09-949-016-2648 |
| 15         | 430.8  | 18.3        | 1763   | 4  | US-09-949-016-2649 |
| 16         | 430.8  | 18.3        | 1788   | 4  | US-09-976-594-313  |
| 17         | 263    | 11.2        | 263    | 4  | US-09-016-434-310  |
| 18         | 229.2  | 9.7         | 913    | 4  | US-09-016-434-712  |
| 19         | 203.4  | 8.6         | 492    | 4  | US-09-621-976-3125 |
| 20         | 185.4  | 7.9         | 479    | 4  | US-09-621-976-3124 |
| 21         | 123.8  | 5.3         | 378    | 1  | US-08-700-575-2    |
| 22         | 95.2   | 4.0         | 2061   | 2  | US-08-835-170-1    |
| 23         | 95.2   | 4.0         | 2061   | 3  | US-09-359-257-1    |
| 24         | 95.2   | 4.0         | 2061   | 3  | US-09-371-674-1    |
| 25         | 95.2   | 4.0         | 2327   | 2  | US-08-835-170-3    |
| 26         | 95.2   | 4.0         | 2327   | 3  | US-09-359-257-3    |
| 27         | 95.2   | 4.0         | 2327   | 3  | US-09-371-674-3    |

|    |      |     |        |   |                      |                    |
|----|------|-----|--------|---|----------------------|--------------------|
| 28 | 89   | 3.8 | 621    | 4 | US-09-248-796A-4385  | Sequence 4385, Ap  |
| 29 | 84.8 | 3.6 | 308    | 4 | US-09-621-976-9631   | Sequence 9631, Ap  |
| 30 | 81.4 | 3.5 | 260    | 4 | US-09-313-294A-2299  | Sequence 2299, Ap  |
| 31 | 75.6 | 3.2 | 3723   | 4 | US-09-949-016-3590   | Sequence 3590, Ap  |
| 32 | 75.6 | 3.2 | 71251  | 4 | US-09-949-016-15332  | Sequence 15332, A  |
| 33 | 68.8 | 2.9 | 282    | 4 | US-09-313-294A-1168  | Sequence 1168, Ap  |
| 34 | 67.2 | 2.9 | 2424   | 4 | US-09-614-221A-518   | Sequence 518, App  |
| 35 | 66.4 | 2.8 | 466    | 4 | US-09-513-999C-11373 | Sequence 11373, A  |
| 36 | 66   | 2.8 | 362    | 4 | US-09-949-016-12776  | Sequence 12776, A  |
| 37 | 63.6 | 2.7 | 187169 | 4 | US-09-949-016-15940  | Sequence 15940, A  |
| 38 | 63.6 | 2.7 | 191569 | 4 | US-09-949-016-14390  | Sequence 14390, A  |
| 39 | 61.8 | 2.6 | 25882  | 4 | US-09-949-016-14391  | Sequence 14391, A  |
| 40 | 61.8 | 2.6 | 25882  | 4 | US-09-806-708B-22    | Sequence 22, Appl  |
| 41 | 60.6 | 2.6 | 1141   | 4 | US-09-949-016-123504 | Sequence 123504, A |
| 42 | 60.6 | 2.6 | 59519  | 4 | US-09-949-016-12378  | Sequence 12378, A  |
| 43 | 60.6 | 2.6 | 119153 | 4 | US-08-487-826B-13    | Sequence 13, Appl  |
| 44 | 60.4 | 2.6 | 19124  | 2 | US-09-949-016-17296  | Sequence 17296, A  |
| 45 | 60.4 | 2.6 | 251672 | 4 | US-09-949-016-17296  | Sequence 17296, A  |

ALIGNMENTS

RESULT 1  
US-09-810-671-1  
; Sequence 1, Application US/09810671  
; Patent No. 6455291  
; GENERAL INFORMATION:  
; APPLICANT: YAN Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CLO00758  
; CURRENT APPLICATION NUMBER: US/09/810,671  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2354  
; TYPE: DNA  
; ORGANISM: Human  
US-09-810-671-1

|                       |                 |   |           |              |
|-----------------------|-----------------|---|-----------|--------------|
| Query Match           | 100.0%;         | Score 2354;   | DB 3;     | Length 2354; |
| Best Local Similarity | 100.0%;         | Pred. No. 0;  |           |              |
| Matches 2354;         | Conservative 0; | Mismatches 0;   | Indels 0; | Gaps 0;      |
| Qy                    | 1               | GCAGCTGGGGTTACTTTTAAAAACATGCTCCATGTGCATCCCTCTTGAAGTTGCGACT    | 60        |              |
| Db                    | 1               | GCAGCTGGGGTTACTTTTAAAAACATGCTCCATGTGCATCCCTCTTGAAGTTGCGACT    | 60        |              |
| Qy                    | 61              | CTGTTGAAGAGGACACTCCAGTCAATTTTAGAGCAAGGTCCTTGAATGAGCGAG        | 120       |              |
| Db                    | 61              | CTGTTGAAGAGGACACTCCAGTCAATTTTAGAGCAAGGTCCTTGAATGAGCGAG        | 120       |              |
| Qy                    | 121             | ATTATCGGACCGGAGATACCTTGACGAATACAGGAATGACTCTGTGAAGGATATGTC     | 180       |              |
| Db                    | 121             | ATTATCGGACCGGAGATACCTTGACGAATACAGGAATGACTCTGTGAAGGATATGTC     | 180       |              |
| Qy                    | 181             | CTAGACATTATCA CAGAGACATTGAAAGCGGGTATCGAATCCACTGCGAGTAATCTTCAG | 240       |              |
| Db                    | 181             | CTAGACATTATCA CAGAGACATTGAAAGCGGGTATCGAATCCACTGCGAGTAATCTTCAG | 240       |              |
| Qy                    | 241             | TCCGACGAGGAGGAGGAGTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGT     | 300       |              |
| Db                    | 241             | TCCGACGAGGAGGAGGAGTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGT     | 300       |              |
| Qy                    | 301             | CACGTTTGAAGAGGACCGAAGGAAAAGATCCAGGAGTATAGAGGATCATGAGGAGGTC    | 360       |              |
| Db                    | 301             | CACGTTTGAAGAGGACCGAAGGAAAAGATCCAGGAGTATAGAGGATCATGAGGAGGTC    | 360       |              |
| Qy                    | 361             | ACCTGATCTGTCAAAGTGGAGACGTTCTTAAGAGCAAGATATCAATCGTGACACTTTGG   | 420       |              |



```
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-1

Query Match      100.0%; Score 2354; DB 4; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGCTGGGTTACTTTAAAAAACATGCTCATGTGATGCTCTTGAAGCTTCGCAC 60
DB 1 GCCAGCTGGGTTACTTTAAAAAACATGCTCATGTGATGCTCTTGAAGCTTCGCAC 60

QY 61 CTGTTGAAGAGACACTCATCCAGTCATTATTAGAAGCAAGTCCCTTGAATGAGCG 120
DB 61 CTGTTGAAGAGACACTCATCCAGTCATTATTAGAAGCAAGTCCCTTGAATGAGCG 120

QY 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTCTGTGAAGGATATG 180
DB 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTCTGTGAAGGATATG 180

QY 181 CTAGACATTATCACAGAGACATTTGAAGCGGTATCGAATCCACATGCGATTAATCT 240
DB 181 CTAGACATTATCACAGAGACATTTGAAGCGGTATCGAATCCACATGCGATTAATCT 240

QY 241 TCCGACGAGGAGAGAGCTCTTAAAGAGCGCAATAGACATGTTCAAGTCATCAGT 300
DB 241 TCCGACGAGGAGAGAGCTCTTAAAGAGCGCAATAGACATGTTCAAGTCATCAGT 300

QY 301 CACGTTTCCAAGAGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGG 360
DB 301 CACGTTTCCAAGAGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGG 360

QY 361 ACCTGATCTGCAAGTGGAGAGCTTTCAAGAGCAAGATATAAATCGTGGACATTTGG 420
DB 361 ACCTGATCTGCAAGTGGAGAGCTTTCAAGAGCAAGATATAAATCGTGGACATTTGG 420

QY 421 GTGAAGGAGCCCTTGGCAAGGTTGTAGAGTGCAATGATGATGATGATGATGATG 480
DB 421 GTGAAGGAGCCCTTGGCAAGGTTGTAGAGTGCAATGATGATGATGATGATGATG 480

QY 481 TAGCAGTGAATAATCGTAAATAATGAGCGCGTTACCGTGAAGCAGCTCGTTTCA 540
DB 481 TAGCAGTGAATAATCGTAAATAATGAGCGCGTTACCGTGAAGCAGCTCGTTTCA 540

QY 541 AAGTATTAGAGACATTAATAAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGAT 600
DB 541 AAGTATTAGAGACATTAATAAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGAT 600

QY 601 TAGAATGTTTGCATCATGTCATGTTGTATTTGTGTTGTTGTTGTTGTTGTTGTT 660
DB 601 TAGAATGTTTGCATCATGTCATGTTGTATTTGTGTTGTTGTTGTTGTTGTTGTT 660

QY 661 CTTACGATTTCAATAAGAAACAGCTTCTGCCATTTCAATTTCAATTCACACATCAG 720
DB 661 CTTACGATTTCAATAAGAAACAGCTTCTGCCATTTCAATTTCAATTCACACATCAG 720

QY 721 TGGCGTATCAGATCTGCCAGTCAATAAATTTTTTATCATATAAATAAATTAACCCAT 780
DB 721 TGGCGTATCAGATCTGCCAGTCAATAAATTTTTTATCATATAAATAAATTAACCCAT 780

QY 781 ATCTGAAGCCTGAAATAATTTGTTTGTGAAGTCTGACATATGATGATGATGATGAT 840
DB 781 ATCTGAAGCCTGAAATAATTTGTTTGTGAAGTCTGACATATGATGATGATGATGAT 840
```

```
QY 841 AAATGAAACGTTGATGAACGCACTGAAACACACAGATATCAAGTTGTTGATTTGAA 900
DB 841 AAATGAAACGTTGATGAACGCACTGAAACACACAGATATCAAGTTGTTGATTTGAA 900

QY 901 GTGCAACGTTATGATGAACATCAAGTACTTTGGTGTCTACCGGCACTACAGAGCT 960
DB 901 GTGCAACGTTATGATGAACATCAAGTACTTTGGTGTCTACCGGCACTACAGAGCT 960

QY 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCGAGCCCTTGTGATGTTTGGAGCATAG 1020
DB 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCGAGCCCTTGTGATGTTTGGAGCATAG 1020

QY 1021 TTCTTATTGAATATTAACCTTGGTTTTCACAGTCTTTTCAGACTCATGATAGTAA 1080
DB 1021 TTCTTATTGAATATTAACCTTGGTTTTCACAGTCTTTTCAGACTCATGATAGTAA 1080

QY 1081 TGGCAATGATGGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAA 1140
DB 1081 TGGCAATGATGGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAA 1140

QY 1141 AACGCAAGTATTTTCCACCATAAACAGCTAGATTGGGATGAACACACAGTTCCT 1200
DB 1141 AACGCAAGTATTTTCCACCATAAACAGCTAGATTGGGATGAACACACAGTTCCT 1200

QY 1201 ATGTTAGAGAGCGCTGCAACCGTTTGAAGGAATTTATGCTTTGTCATGATGA 1260
DB 1201 ATGTTAGAGAGCGCTGCAACCGTTTGAAGGAATTTATGCTTTGTCATGATGA 1260

QY 1261 AGAAACTGTTTGAACCTGTCAGCATCTCTTTTGAATTTAAAGAAAGAAATGGA 1320
DB 1261 AGAAACTGTTTGAACCTGTCAGCATCTCTTTTGAATTTAAAGAAAGAAATGGA 1320

QY 1321 TGGATGAAGAGCTGTCAGCATCTCTTTTGAATTTAAAGAAAGAAATGGAATG 1380
DB 1321 TGGATGAAGAGCTGTCAGCATCTCTTTTGAATTTAAAGAAAGAAATGGAATG 1380

QY 1381 AGTGTCTTACTATATATCTCTTAGAAGAGATTAATTAAGACTGTGTGAGTCA 1440
DB 1381 AGTGTCTTACTATATATCTCTTAGAAGAGATTAATTAAGACTGTGTGAGTCA 1440

QY 1441 CATTCATATTTTGTGTAACATTAATTTTGTACAGTTAAGTGAATATTTGATG 1500
DB 1441 CATTCATATTTTGTGTAACATTAATTTTGTACAGTTAAGTGAATATTTGATG 1500

QY 1501 TTTTGTATCAATAGCATAAATTAATCTTTTAAAGAGATGTTGTTGATTAATCA 1560
DB 1501 TTTTGTATCAATAGCATAAATTAATCTTTTAAAGAGATGTTGTTGATTAATCA 1560

QY 1561 AAATTAATAATTAATTTTCTTTTGAATTTACCATTCTTTTAAATATCTTTT 1620
DB 1561 AAATTAATAATTAATTTTCTTTTGAATTTACCATTCTTTTAAATATCTTTT 1620

QY 1621 TGTGTCCAGTGAATAATGATGATCTTGTGCTTTGTACATGAGAGTCACTCTGA 1680
DB 1621 TGTGTCCAGTGAATAATGATGATCTTGTGCTTTGTACATGAGAGTCACTCTGA 1680

QY 1681 GATTTTCTTCAAGTAAAGGAAATCTTTCAGTACTTTTATTTTAAAGGAAATTT 1740
DB 1681 GATTTTCTTCAAGTAAAGGAAATCTTTCAGTACTTTTATTTTAAAGGAAATTT 1740

QY 1741 TATACCTTCAAAATTTAGAACTTAACTTTTAAAGTCTTTTCTTCTGTAATTT 1800
DB 1741 TATACCTTCAAAATTTAGAACTTAACTTTTAAAGTCTTTTCTTCTGTAATTT 1800

QY 1801 ATTATTTAATCTCTAGATAAGCGGTACTAGAAACCAAACTCAGAAAAATGTT 1860
DB 1801 ATTATTTAATCTCTAGATAAGCGGTACTAGAAACCAAACTCAGAAAAATGTT 1860

QY 1861 TAGAATTTCTAATAATTTTAAAGTGTGTTGTTTCTTTTCAATTTGGGTTGAT 1920
DB 1861 TAGAATTTCTAATAATTTTAAAGTGTGTTGTTTCTTTTCAATTTGGGTTGAT 1920

QY 1921 ACCAGACATTCATGGAAGGCATGTCAGTTTGTCCATTTGTGACAGTTTGT 1980
```

```

Db 1921 ACCAGACATTCATGGAAGGCGATGCGAGTTGTGTCATTCGACAGTTTCTTTAATAAACC 1980
QY 1981 ACATACACACTTTATTTAAGATTAAATCTAAGTGAAGTCAAGTCTGGAAGTGGACAT 2040
Db 1981 ACATACACACTTTATTTAAGATTAAATCTAAGTGAAGTCAAGTCTGGAAGTGGACAT 2040
QY 2041 TTCCAAAGTATGTTGGTGAAGTCAAGATATATAAATAGAAATCTTGATGAGAGGTTTCAG 2100
Db 2041 TTCCAAAGTATGTTGGTGAAGTCAAGATATATAAATAGAAATCTTGATGAGAGGTTTCAG 2100
QY 2101 TTTTAAATACCAAGTCTTTAGAGTCTTAAATGAGTGGCAGCATCTGTTTATCAAAATGACA 2160
Db 2101 TTTTAAATACCAAGTCTTTAGAGTCTTAAATGAGTGGCAGCATCTGTTTATCAAAATGACA 2160
QY 2161 TAAATACCTAACCCTATTAAGATTAAAGTTTAAATAGGCAATTTATGCTGTGATAAT 2220
Db 2161 TAAATACCTAACCCTATTAAGATTAAAGTTTAAATAGGCAATTTATGCTGTGATAAT 2220
QY 2221 TCTTACGGGAGAAAGAGGATTTGATTGGAAGCAGTTTGGGAAGAGTGTGCTGTAAT 2280
Db 2221 TCTTACGGGAGAAAGAGGATTTGATTGGAAGCAGTTTGGGAAGAGTGTGCTGTAAT 2280
QY 2281 TTCCAGAAATTAATGATTGTTACATAAACTTTTGTGACTTCAGAAAAAATAAATAA 2340
Db 2281 TTCCAGAAATTAATGATTGTTACATAAACTTTTGTGACTTCAGAAAAAATAAATAA 2340
QY 2341 AACAAAAAATAAAC 2354
Db 2341 AACAAAAAATAAAC 2354

```

```

RESULT 3
US-10-339-656-1
; Sequence 1, Application US/10339656
; Patent No. 673978
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-1

```

```

Query Match 100.0%; Score 2354; DB 4; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGCTGGGTTACTTTAAATAAATGCTCCATGTCATCCCTCTTGAAGCTTCGCACT 60
Db 1 GCCAGCTGGGTTACTTTAAATAAATGCTCCATGTCATCCCTCTTGAAGCTTCGCACT 60
QY 61 CTGTTGAAGAGACACTCATCCAGTCAATTTTGAAGCAAGGTCCTTGAATGAGCGAG 120
Db 61 CTGTTGAAGAGACACTCATCCAGTCAATTTTGAAGCAAGGTCCTTGAATGAGCGAG 120
QY 121 ATTATCGGACCGGAGATACGTTGACGATACAGGAATGACTACTGTGAAGATATGTC 180
Db 121 ATTATCGGACCGGAGATACGTTGACGATACAGGAATGACTACTGTGAAGATATGTC 180

```

```

QY 181 CTAGACATTTATCAGAGACATTTGAAAGCGGTTATCGAATCCACTGCAAGTAAATCTTCAG 240
Db 181 CTAGACATTTATCAGAGACATTTGAAAGCGGTTATCGAATCCACTGCAAGTAAATCTTCAG 240
QY 241 TCCGAGAGAGAGAGAGAGTCTTAAAGGAAGCCCAATAGACACTGTTCAAGTCAATCAGT 300
Db 241 TCCGAGAGAGAGAGAGTCTTAAAGGAAGCCCAATAGACACTGTTTCAAGTCAATCAGT 300
QY 301 CACGTTCCGAGAGAGAGAGATCCGAGAGGAAAGATCCAGAGGTATAGAGATGATGAGAGGCTC 360
Db 301 CACGTTCCGAGAGAGAGAGATCCGAGAGGAAAGATCCAGAGGTATAGAGATGATGAGAGGCTC 360
QY 361 ACCTGATCTGTCGAAAGTGGAGACGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGG 420
Db 361 ACCTGATCTGTCGAAAGTGGAGACGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGG 420
QY 421 GTGAAGAGAGCCTTTGGCAAGTGTAGAGTGCATTTGATGCGCATGATGCGCATGCGCATG 480
Db 421 GTGAAGAGAGCCTTTGGCAAGTGTAGAGTGCATTTGATGCGCATGATGCGCATGCGCATG 480
QY 481 TAGCAGTGAAATCGTAAATAAGTGTAGGCGGTTACCGTGAAGCAGCTCGTTTCAGAAATCC 540
Db 481 TAGCAGTGAAATCGTAAATAAGTGTAGGCGGTTACCGTGAAGCAGCTCGTTTCAGAAATCC 540
QY 541 AAGTATTAGAGCACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGTC 600
Db 541 AAGTATTAGAGCACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGTC 600
QY 601 TAGAATGGTTTGTGATCATGCTGATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 660
Db 601 TAGAATGGTTTGTGATCATGCTGATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 660
QY 661 CTTACGATTTCAATTAAGAGAAACAGCTTTCTGCAATTTCAAAATGACCAATCAAGGAGCA 720
Db 661 CTTACGATTTCAATTAAGAGAAACAGCTTTCTGCAATTTCAAAATGACCAATCAAGGAGCA 720
QY 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTACATCAATAAATAAATTAACCCATACAG 780
Db 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTACATCAATAAATAAATTAACCCATACAG 780
QY 781 ATCTGAAGCCTGAAATAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
Db 781 ATCTGAAGCCTGAAATAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
QY 841 AAATGAACGTTGATGAACGACACTGAAACACAGATATCAAAAGTGTGTTGTTGTTGTTGTT 900
Db 841 AAATGAACGTTGATGAACGACACTGAAACACAGATATCAAAAGTGTGTTGTTGTTGTTGTT 900
QY 901 GTGCAACGTTGATGATGAACATCAAGTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960
Db 901 GTGCAACGTTGATGATGAACATCAAGTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960
QY 961 CCGAGGTCATTTGGCTTTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1020
Db 961 CCGAGGTCATTTGGCTTTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1020
QY 1021 TTTCTATTGAATATTTACCTTGGTTTTCAGAGTCTTTTACAGCTCATGATAGTAAAGAGCACC 1080
Db 1021 TTTCTATTGAATATTTACCTTGGTTTTCAGAGTCTTTTACAGCTCATGATAGTAAAGAGCACC 1080
QY 1081 TGGCAATGATGAACGAATATTAGGACCCATACCAACACATGATTCAGAAAAACAAGAA 1140
Db 1081 TGGCAATGATGAACGAATATTAGGACCCATACCAACACATGATTCAGAAAAACAAGAA 1140
QY 1141 AACGCAAGTATTTTCAACCAATCCAGTGTAGATGGGATGAACACAGTCTCTGTTAGTAT 1200
Db 1141 AACGCAAGTATTTTCAACCAATCCAGTGTAGATGGGATGAACACAGTCTCTGTTAGTAT 1200
QY 1201 ATGTTAGGAGAGCGTGCAGAAACCGTTGAAGGAATTTATGCTTTGTTCATGATGAAGAACATG 1260
Db 1201 ATGTTAGGAGAGCGTGCAGAAACCGTTGAAGGAATTTATGCTTTGTTCATGATGAAGAACATG 1260

```

QY 1261 AGAACTCTTTGACCTGTTGCGAAGATGTTAGAAATATGATCCAACTCAAGAAATTACCT 1320  
Db 1261 AGAACTCTTTGACCTGTTGCGAAGATGTTAGAAATATGATCCAACTCAAGAAATTACCT 1320  
QY 1321 TGGATGAAGCATTCGAGCATCTTTCTTTGACTTATTAAGAAAGAAATGAAATCGGAATC 1380  
Db 1321 TGGATGAAGCATTCGAGCATCTTTCTTTGACTTATTAAGAAAGAAATGAAATCGGAATC 1380  
QY 1381 AGTGTCTTACTATATCTTCTAGAGAGAGATTACTTAAGACGTGTGTCACTCAACTAAA 1440  
Db 1381 AGTGTCTTACTATATCTTCTAGAGAGAGATTACTTAAGACGTGTGTCACTCAACTAAA 1440  
QY 1441 CATCTCAATATTTTGTAAACATTAATATTTGTACAGTAAGTGAATATTTGTATG 1500  
Db 1441 CATCTCAATATTTTGTAAACATTAATATTTGTACAGTAAGTGAATATTTGTATG 1500  
QY 1501 TTTTGTATCAATAGCATTAATTAACCTTTGTAAGCAAGTATGCTTTGATAATGCATTAGAA 1560  
Db 1501 TTTTGTATCAATAGCATTAATTAACCTTTGTAAGCAAGTATGCTTTGATAATGCATTAGAA 1560  
QY 1561 AAATTAATAATTAATTTTCTTTTGAATTAACATTTTAAATACCTTTTGAATATCTCT 1620  
Db 1561 AAATTAATAATTAATTTTCTTTTGAATTAACATTTTAAATACCTTTTGAATATCTCT 1620  
QY 1621 TGTGTCCAGTGAATAATGTGATGATCTTGCTTTGTACATGAGGTCACCTCTGAAAT 1680  
Db 1621 TGTGTCCAGTGAATAATGTGATGATCTTGCTTTGTACATGAGGTCACCTCTGAAAT 1680  
QY 1681 GATTTTTTTTTCAGTAAAGGAAATCTTCACACTTATATATTTCTTAAAGGAATATCTTTA 1740  
Db 1681 GATTTTTTTTTCAGTAAAGGAAATCTTCACACTTATATATTTCTTAAAGGAATATCTTTA 1740  
QY 1741 TATACTTCAAAATTTAGAACTTAACTTTAAAGTTTCTTCTGTAATTTGTAACGGGTG 1800  
Db 1741 TATACTTCAAAATTTAGAACTTAACTTTAAAGTTTCTTCTGTAATTTGTAACGGGTG 1800  
QY 1801 ATTAATTAATCTCTAGATAAGCAGTACTAGAAACCAAACTCAGAAAAATGTTTACTGT 1860  
Db 1801 ATTAATTAATCTCTAGATAAGCAGTACTAGAAACCAAACTCAGAAAAATGTTTACTGT 1860  
QY 1861 TAGAATCTATTAATTTTAAAGTGTGTAATTTCTTTTCAATTTGGGTGATGTCAGGTGATA 1920  
Db 1861 TAGAATCTATTAATTTTAAAGTGTGTAATTTCTTTTCAATTTGGGTGATGTCAGGTGATA 1920  
QY 1921 ACCAGACATTCATGAAAGGATGAGTTTGTCCATTTGTGACAGTTTCTTTTAAATAAAC 1980  
Db 1921 ACCAGACATTCATGAAAGGATGAGTTTGTCCATTTGTGACAGTTTCTTTTAAATAAAC 1980  
QY 1981 ACATACACACTTTATTTAAGATTAATAATCTAACTGGAAGTCAGTTTGGAAAAATGGACAT 2040  
Db 1981 ACATACACACTTTATTTAAGATTAATAATCTAACTGGAAGTCAGTTTGGAAAAATGGACAT 2040  
QY 2041 TTCCAGTATGTTGGTGCAGTCACAGATATAAATAAGAAATTTCTGATGAGAGTTTCAG 2100  
Db 2041 TTCCAGTATGTTGGTGCAGTCACAGATATAAATAAGAAATTTCTGATGAGAGTTTCAG 2100  
QY 2101 TTTTAAATACCAAGTCCTTAGAGCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2160  
Db 2101 TTTTAAATACCAAGTCCTTAGAGCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2160  
QY 2161 TAAATACGTAACCTTATTAAGATTAAGTTTATTAATTAAGGCAATTTATGCTGTGATAAT 2220  
Db 2161 TAAATACGTAACCTTATTAAGATTAAGTTTATTAATTAAGGCAATTTATGCTGTGATAAT 2220  
QY 2221 TCTTACGGGAGAAAGAGATTTGATTTGAAAGCAGTTTGGGAAGAAAGTGTCTGTAAT 2280  
Db 2221 TCTTACGGGAGAAAGAGATTTGATTTGAAAGCAGTTTGGGAAGAAAGTGTCTGTAAT 2280  
QY 2281 TTCCAGAAATTTAATTTGATTTGATACATAAACTTTTGTGACTTTCAGAAAAAATAAAAA 2340  
Db 2281 TTCCAGAAATTTAATTTGATTTGATACATAAACTTTTGTGACTTTCAGAAAAAATAAAAA 2340  
QY 2341 AACAAAAAATAAAC 2354

Db 2341 AACAAAAAATAAAC 2354

## RESULT 4

US-09-016-000-9  
; Sequence 9, Application US/09016000  
; Patent No. 5962232  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Bandman, Olga  
; APPLICANT: Akerbom, Ingrid B.  
; APPLICANT: Shah, Purvi  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,000  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0465 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2446 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: NEUTFWT01  
; CLONE: 339963  
US-09-016-000-9

Query Match 87.2%; Score 2052.6; DB 2; Length 2446;

Best Local Similarity 95.8%; Pred. No. 0;  
Matches 2155; Conservative 0; Mismatches 4; Indels 90; Gaps 1;

QY 85 GTCATTATTTAGAACGAAGTCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG 144  
Db 288 GTCATTATTTAGAACGAAGTCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG 347  
QY 145 ACGAATACAGGAATGACTACTGTGAAGTATGTTCTAGACATTATCACAGACATTG 204  
Db 348 ACGAATACAGGAATGACTACTGTGAAGTATGTTCTAGACATTATCACAGACATTG 407  
QY 205 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGCGAGGAGAGAGCAGTCTCTA 264  
Db 408 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGCGAGGAGAGAGCAGTCTCTA 467  
QY 265 AAAGGAAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAAAGAGCCACCGAAGGA 324



|    |      |  |      |
|----|------|--|------|
| Db | 468  | AAAGGAAGCGCAATAGACACATGTTCAAGTCATCAGTCACGTTCCGATGA-----          | 515  |
| Qy | 325  | AAAGATCCAGGATATAGAGGATGATGAGGAGTCACTGATCTGTCAAAGTGGAGACG         | 384  |
| Db | 516  | -----  | 515  |
| Qy | 385  | TTCTAAGACCAAGATATGAATTCGTGGACACATTTGGGTGAAGGAGCCCTTGGCAAAAGTTG   | 444  |
| Db | 516  | -----AATCCGTGGACACATTTGGGTGAAGGAGCCCTTGGCAAAAGTTG                | 557  |
| Qy | 445  | TAGAGTGCAATGATCATGGCATGGATGCGATGTCATGTAGCAGTGAATAATCGTAAATAATG   | 504  |
| Db | 558  | TAGAGTGCAATGATCATGGCATGGATGCGATGTCATGTAGCAGTGAATAATCGTAAATAATG   | 617  |
| Qy | 505  | TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAGTATTAAGAGCAGTAAATAGTA      | 564  |
| Db | 618  | TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAGTATTAAGAGCAGTAAATAGTA      | 677  |
| Qy | 565  | CTGATCCCAATAGTGTCTTCCGATGTCTCCAGATGCTAGAAATGGTTTGATCATCATGGTC    | 624  |
| Db | 678  | CTGATCCCAATAGTGTCTTCCGATGTCTCCAGATGCTAGAAATGGTTTGATCATCATGGTC    | 737  |
| Qy | 635  | ATGTTTGTATTGTTGAACTACTGGAACCTAGTACTTACGATTTCAATTAAGAAACA         | 684  |
| Db | 738  | ATGTTTGTATTGTTGAACTACTGGAACCTAGTACTTACGATTTCAATTAAGAAACA         | 797  |
| Qy | 685  | GCTTCTGCCATTTCAAAATGACCAATCAGCAGATGGCGTATCAGATCTGCCAGTCAA        | 744  |
| Db | 798  | GCTTCTGCCATTTCAAAATGACCAATCAGCAGATGGCGTATCAGATCTGCCAGTCAA        | 857  |
| Qy | 745  | TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAAGCTCGAAATATTTTGT   | 804  |
| Db | 858  | TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAAGCTCGAAATATTTTGT   | 917  |
| Qy | 805  | TTGTGAAGTCTGACTATGATGATCAATATAATTAATTAATTAATTAATTAATTAATTAATTA   | 864  |
| Db | 918  | TTGTGAAGTCTGACTATGATGATCAATATAATTAATTAATTAATTAATTAATTAATTAATTA   | 977  |
| Qy | 865  | TGAAAAACACAGATATCAAGTTGTTGACTTTGGAAGTGCACGATGATGATGAACATC        | 924  |
| Db | 978  | TGAAAAACACAGATATCAAGTTGTTGACTTTGGAAGTGCACGATGATGATGAACATC        | 1037 |
| Qy | 925  | ACAGTACTTTGTTGTTCTACCCGGCACTACAGAGTCTCCGAGTCAATTTTGGCTTTAGGTT    | 984  |
| Db | 1038 | ACAGTACTTTGTTGTTCTACCCGGCACTACAGAGTCTCCGAGTCAATTTTGGCTTTAGGTT    | 1097 |
| Qy | 985  | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTAATTAATTAATTAATTAAT  | 1044 |
| Db | 1098 | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTAATTAATTAATTAATTAAT  | 1157 |
| Qy | 1045 | TCACAGTCTTTTCAGACTCATGATGATTAAGAGCCTCGCATGATGGAACGAAATATTAG      | 1104 |
| Db | 1158 | TCACAGTCTTTTCAGACTCATGATGATTAAGAGCCTCGCATGATGGAACGAAATATTAG      | 1217 |
| Qy | 1105 | GAACCATACCAACACACATGATTTCAGAAAAACGAAAGTATTTTCAACATAACC           | 1164 |
| Db | 1218 | GAACCATACCAACACACATGATTTCAGAAAAACGAAAGTATTTTCAACATAACC           | 1277 |
| Qy | 1165 | AGCTAGATTGGATGAACACACAGTCTCTGTTGATGATGTTAGGAGAGCGTGCAAAACCGT     | 1224 |
| Db | 1278 | AGCTAGATTGGATGAACACACAGTCTCTGTTGATGATGTTAGGAGAGCGTGCAAAACCGT     | 1337 |
| Qy | 1225 | TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAAATGTTTCCCTGTTCCGAA        | 1284 |
| Db | 1338 | TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAAATGTTTCCCTGTTCCGAA        | 1397 |
| Qy | 1285 | GAAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCTTT   | 1344 |
| Db | 1398 | GAAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCTTT   | 1457 |
| Qy | 1345 | TCTTTGACTTATTAAGAAATGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT | 1404 |

|    |      |   |      |
|----|------|---|------|
| Db | 1458 | TCCTTGACTTATTAAGAAATGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT    | 1517 |
| Qy | 1405 | AGAAGAGATTACTTAAGAGTGTGTGAGTCACTAAACATTAATTAATTAATTAATTAATTAATTAAT  | 1464 |
| Db | 1518 | AGAAGAGATTACTTAAGAGTGTGTGAGTCACTAAACATTAATTAATTAATTAATTAATTAATTAAT  | 1577 |
| Qy | 1465 | AAATTAATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATTAATTAATTAATTAATTAAT  | 1524 |
| Db | 1578 | AAATTAATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATTAATTAATTAATTAATTAAT  | 1637 |
| Qy | 1525 | TTGTTAAGCAAGTATGTTTGTATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAAT   | 1584 |
| Db | 1638 | TTGTTAAGCAAGTATGTTTGTATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAAT   | 1697 |
| Qy | 1585 | GAAATTAACATTTTAAATACCTTTGAAATATCTTTGTGTCCAGTGAATAATGATGATG          | 1644 |
| Db | 1698 | GAAATTAACATTTTAAATACCTTTGAAATATCTTTGTGTCCAGTGAATAATGATGATG          | 1757 |
| Qy | 1645 | ATCTTGCCCTTTGTACATGGAGTCACTCTGAAAGTGAATTTTGTAGTAAAGGAAAT            | 1704 |
| Db | 1758 | ATCTTGCCCTTTGTACATGGAGTCACTCTGAAAGTGAATTTTGTAGTAAAGGAAAT            | 1817 |
| Qy | 1705 | CTTGACTACTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAAC      | 1764 |
| Db | 1818 | CTTGACTACTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAAC      | 1877 |
| Qy | 1765 | TTTAAAGTCTTTCTCTGTAAATTTGTAACGGGTGATTATTAATTAATTAATTAATTAATTAATTAAT | 1824 |
| Db | 1878 | TTTAAAGTCTTTCTCTGTAAATTTGTAACGGGTGATTATTAATTAATTAATTAATTAATTAATTAAT | 1937 |
| Qy | 1825 | GGTACTAGAAACCAAACTCAGAAAACTCTGTTAGAAATTTCTATTAATTAATTAATTAATTAAT    | 1884 |
| Db | 1938 | GGTACTAGAAACCAAACTCAGAAAACTCTGTTAGAAATTTCTATTAATTAATTAATTAATTAAT    | 1997 |
| Qy | 1885 | TTGTATTTCTTTTCTTCTGTTGATGTCAGGGTGAATTAACCAAGTGAATTAATTAATTAATTAAT   | 1944 |
| Db | 1998 | TTGTATTTCTTTTCTTCTGTTGATGTCAGGGTGAATTAACCAAGTGAATTAATTAATTAATTAAT   | 2057 |
| Qy | 1945 | CAGTTTGTCCACTGTGACAGTTTGTAAATAAAACCAATACACATCACTTTTATTAAGATTA       | 2004 |
| Db | 2058 | CAGTTTGTCCACTGTGACAGTTTGTAAATAAAACCAATACACATCACTTTTATTAAGATTA       | 2117 |
| Qy | 2005 | AAATCTAACTGGAAAGTCACTTGGAAATGGAATTCATTTCCAAAGTATGTTTGGTGAAGTCAAC    | 2064 |
| Db | 2118 | AAATCTAACTGGAAAGTCACTTGGAAATGGAATTCATTTCCAAAGTATGTTTGGTGAAGTCAAC    | 2177 |
| Qy | 2065 | AGATATAAAATAGAAATTTCTGTAGAGAGTTCAGTTTAAATTAACCAAGTCTTTAGGAG         | 2124 |
| Db | 2178 | AGATATAAAATAGAAATTTCTGTAGAGAGTTCAGTTTAAATTAACCAAGTCTTTAGGAG         | 2237 |
| Qy | 2125 | TCCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAAGAAAT      | 2184 |
| Db | 2238 | TCCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAAGAAAT      | 2297 |
| Qy | 2185 | AAGTTTATTAATTAAGCAATTTATGTCGTGATTAATTTCTTACGGGAGAAAGGATTTGA         | 2244 |
| Db | 2298 | AAGTTTATTAATTAAGCAATTTATGTCGTGATTAATTTCTTACGGGAGAAAGGATTTGA         | 2357 |
| Qy | 2245 | TTGGAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAATTAATTAATTAATTAAT | 2304 |
| Db | 2358 | TTGGAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAATTAATTAATTAATTAAT | 2417 |
| Qy | 2305 | CATAAATCTTTTGTGACTTCAGAAAAAATAA 2333                                |      |
| Db | 2418 | CATAAATCTTTTGTGACTTCAGAAAAAATAA 2446                                |      |

RESULT 5  
 US-09-023-655-699  
 ; Sequence 699, Application US/09023655  
 ; Patent No 6607879  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cocks, Benjamin G.



APPLICANT: Susan G. Stuart  
 APPLICANT: Jeffrey J. Seilhamer  
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
 TITLE OF INVENTION: EXPRESSION  
 NUMBER OF SEQUENCES: 1508  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/023,655  
 FILING DATE: HEREWITH  
 CLASSIFICATION:  
 PRIOR APPLICATION NUMBER:  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0001 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 699:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1456 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: HNT2AGT01  
 CLONE: 488942  
 US-09-023-655-699

Query Match 50.1%; Score 1178.2; DB 4; Length 1456;  
 Best Local Similarity 99.7%; Pred. No. 4.7e-286;  
 Matches 1191; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

|    |      |   |      |
|----|------|---|------|
| QY | 1003 | TTTGAGCATAGGTTGCATCTCTATTGAATATTACCTTGGTTTCACAGTCTTTCAGACTC   | 1062 |
| DB | 262  | TGTAGCGCATAGGTTGCATCTCTATTGAATATTACCTTGGTTTCACAGTCTTTCAGACTC  | 321  |
| QY | 1063 | ATGATAGTAAAGAGCACCTGGCAATGATGGACGAATATTAGGACCCCATACCACACACA   | 1122 |
| DB | 322  | ATGATAGTAAAGAGCACCTGGCAATGATGGACGAATATTAGGACCCCATACCACACACA   | 381  |
| QY | 1123 | TGATTCAGAAAACAAGAAAACGCAAGTATTTTCCACATAACGAGCTAGATTGGGATGAAC  | 1182 |
| DB | 382  | TGATTCAGAAAACAAGAAAACGCAAGTATTTTCCACATAACGAGCTAGATTGGGATGAAC  | 441  |
| QY | 1183 | ACAGTTCTGCTGGTAGATATGTTAGAGACGCTGCAAAACCGTTGAAGGAATTTATGCTTT  | 1242 |
| DB | 442  | ACAGTTCTGCTGGTAGATATGTTAGAGACGCTGCAAAACCGTTGAAGGAATTTATGCTTT  | 501  |
| QY | 1243 | GTCATGATGAAGAACATGAGAAACGTTTGACCTGGTTGCAAGAAATGTTAGAAATATGATC | 1302 |
| DB | 502  | GTCATGATGAAGAACATGAGAAACGTTTGACCTGGTTGCAAGAAATGTTAGAAATATGATC | 561  |
| QY | 1303 | CAACTCAAGAAATACCTTGGATGAAGCAATTCAGCATCCTTTCTTTCGACTTATTAAAAA  | 1362 |
| DB | 562  | CAACTCAAGAAATACCTTGGATGAAGCAATTCAGCATCCTTTCTTTCGACTTATTAAAAA  | 621  |
| QY | 1363 | AGAAATGAATGGGAATCAGTGGTCTTACTATATATCTCTCTAGAGAGATTACTTAAGA    | 1422 |

|    |      |  |      |
|----|------|--|------|
| DB | 622  | AGAAATGAATGGGAATCAGTGGTCTTACTATATATCTCTCTAGAGAGATTACTTAAGA   | 681  |
| QY | 1423 | CTGTGTCAGTCAACTAAACATTTTGTGTAATATTTTGTGTAATATTTTGTACAGTT     | 1482 |
| DB | 682  | CTGTGTCAGTCAACTAAACATTTTGTGTAATATTTTGTGTAATATTTTGTACAGTT     | 741  |
| QY | 1483 | AAAGTAAATATTTGTTTGTATCAATAGCAATTAATTAATCTGTTAAAGCAAGTATGTT   | 1542 |
| DB | 742  | AAAGTAAATATTTGTTTGTATCAATAGCAATTAATTAATCTGTTAAAGCAAGTATGTT   | 801  |
| QY | 1543 | CTTGATATGCAATTAGAAAAATTAATTTTCTTTTGAATTA-CCATTTTAA           | 1601 |
| DB | 802  | CTTGATATGCAATTAGAAAAATTAATTTTCTTTTGAATTA-CCATTTTAA           | 861  |
| QY | 1602 | ATACCTTTGAAATATCTTTGTCAGTGAATAATGATGATCTTGGCTTTTGTACA        | 1661 |
| DB | 862  | ATACCTTTGAAATATCTTTGTCAGTGAATAATGATGATCTTGGCTTTTGTACA        | 921  |
| QY | 1662 | TGGAGGTCACTCTGAAAGTATTTTGTGTAAGGAATCTTGACTACTTTATATT         | 1721 |
| DB | 922  | TGGAGGTCACTCTGAAAGTATTTTGTGTAAGGAATCTTGACTACTTTATATT         | 981  |
| QY | 1722 | CTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAACTTTTAAAGTTTCTTC | 1781 |
| DB | 982  | CTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAACTTTTAAAGTTTCTTC | 1041 |
| QY | 1782 | TGTAATTGTTGAACGGGTGATTATTTAATTAACCTAGATAAGCAGTACTAGAAACCAAAA | 1841 |
| DB | 1042 | TGTAATTGTTGAACGGGTGATTATTTAATTAACCTAGATAAGCAGTACTAGAAACCAAAA | 1101 |
| QY | 1842 | CTCAGAAATGTTTACTGTTAGAATCTTATAAATTTTAAAGTTTGTATCTTTTTCATT    | 1901 |
| DB | 1102 | CTCAGAAATGTTTACTGTTAGAATCTTATAAATTTTAAAGTTTGTATCTTTTTCATT    | 1161 |
| QY | 1902 | GSETGATGTCAGGTGATAACAGACATTCATGGAAGGCGATGAGTTTGTCCATTGTA     | 1961 |
| DB | 1162 | GSETGATGTCAGGTGATAACAGACATTCATGGAAGGCGATGAGTTTGTCCATTGTA     | 1221 |
| QY | 1962 | CAGTTTGTAAATAAAACCATACACATCTTTTAAAGATTAAATCTAACTGGAAGT       | 2021 |
| DB | 1222 | CAGTTTGTAAATAAAACCATACACATCTTTTAAAGATTAAATCTAACTGGAAGT       | 1281 |
| QY | 2022 | CAGTTTGTAAATAAAACCATACACATCTTTTAAAGATTAAATCTAACTGGAAGT       | 2081 |
| DB | 1282 | CAGTTTGTAAATAAAACCATACACATCTTTTAAAGATTAAATCTAACTGGAAGT       | 1341 |
| QY | 2082 | TTCTGATCAGAGGTTTTCAGTTTAAATACCAAGTCTTTAGGAGTCTTAACTTGGCCAGC  | 2141 |
| DB | 1342 | TTCTGATGAGAGGTTTTCAGTTTAAATACCAAGTCTTTAGGAGTCTTAACTTGGCCAGC  | 1401 |
| QY | 2142 | ATCTGTTTATCAAAATGACATAAATACGTAACCTATAAGAAATTAAGTTTATTAAT     | 2196 |
| DB | 1402 | ATCTGTTTATCAAAATGACATAAATACGTAACCTATAAGAAATTAAGTTTATTAAT     | 1456 |

## RESULT 6

US-09-905-999-26  
 ; Sequence 26, Application US/09905999  
 ; Patent No. 6797513  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ULLRICH, Axel  
 ; APPLICANT: NAYLER, Oliver  
 ; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS  
 ; FILE REFERENCE: 038602/0431  
 ; CURRENT APPLICATION NUMBER: US/09/905,999  
 ; CURRENT FILING DATE: 2001-07-17  
 ; PRIOR APPLICATION NUMBER: 09/127,248  
 ; PRIOR FILING DATE: 1999-07-31  
 ; PRIOR APPLICATION NUMBER: PCT/IB97/00946  
 ; PRIOR FILING DATE: 1997-06-17  
 ; PRIOR APPLICATION NUMBER: US 08/877,150  
 ; PRIOR FILING DATE: 1997-06-17  
 ; PRIOR APPLICATION NUMBER: US 60/034,286



QY 1528 TTAAGCAAGTATGGCTTGTGATGATGCAATAGAAAAATTAATAATTAATTTCTTTTGGAA 1587  
Db 18438 TTAAGCAAGTATGGCTTGTGATGATGCAATAGAAAAATTAATAATTAATTTCTTTTGGAA 18497  
QY 1588 ATTACCAATTTTAAATACCTTTGAAATATCCCTTGTGTCAGTGATAAATGATGATGATC 1647  
Db 18498 ATTACCAATTTTAAATACCTTTGAAATATCCCTTGTGTCAGTGATAAATGATGATGATC 18557  
QY 1648 TTGCTTTTGTACATGGAGGTCACTCTGAAAGTGAATTTTGTAGTAAAGGAATCTT 1707  
Db 18558 TTGCTTTTGTACATGGAGGTCACTCTGAAAGTGAATTTTGTAGTAAAGGAATCTT 18617  
QY 1708 GACTACTTTATATCTTAAAGGAATATCTTTATATATCTTAAATTTAGAACTTTA 1767  
Db 18618 GACTACTTTATATCTTAAAGGAATATCTTTATATATCTTAAATTTAGAACTTTA 18677  
QY 1768 AAAAGTTTTCTTCTGTAATTTCTCAAGGGTGATTTATTTAATCTCTAGATAAGCAGGT 1827  
Db 18678 AAAAGTTTTCTTCTGTAATTTCTCAAGGGTGATTTATTTAATCTCTAGATAAGCAGGT 18737  
QY 1828 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGTGTG 1887  
Db 18738 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGTGTG 18797  
QY 1888 TATTCCTTTTCTATTCGGTGATGTCAGGGTGATAACACAGACATTCATGGAAGGCATGCAG 1947  
Db 18798 TATTCCTTTTCTATTCGGTGATGTCAGGGTGATAACACAGACATTCATGGAAGGCATGCAG 18857  
QY 1948 TTTGTCCATTTGACAGTTTGTAAATTAATAAACACACATACACATTTTAAAGATTAATA 2007  
Db 18858 TTTGTCCATTTGACAGTTTGTAAATTAATAAACACACATACACATTTTAAAGATTAATA 18917  
QY 2008 TCTAACTGGAAGTCACTTGGAAAAATGACATTTTCCAAATGATGTTGGTGAGTCACAGA 2067  
Db 18918 TCTAACTGGAAGTCACTTGGAAAAATGACATTTTCCAAATGATGTTGGTGAGTCACAGA 18977  
QY 2068 TATAAAATAGAAATTTCTGATGAGAGGTTTCCAGTTTTTAAATACCAGTCTTTAGAGTCT 2127  
Db 18978 TATAAAATAGAAATTTCTGATGAGAGGTTTCCAGTTTTTAAATACCAGTCTTTAGAGTCT 19037  
QY 2128 TAACTTTGCGAGCATCTGTTTATCAATGACATAAATACGTAACCTTAAGAATTAAG 2187  
Db 19038 TAACTTTGCGAGCATCTGTTTATCAATGACATAAATACGTAACCTTAAGAATTAAG 19097  
QY 2188 TTTATTAATAGGCAATTTATGTCGTGATAATCTTACGGGAGAGAGGATTTGATG 2247  
Db 19098 TTTATTAATAGGCAATTTATGTCGTGATAATCTTACGGGAGAGAGGATTTGATG 19157  
QY 2248 GAAAGCAGTTTGGGAGAAAGTGTGCTGCTGAAATTTCCAGAAATTTAATTTGATGTTACAT 2307  
Db 19158 GAAAGCAGTTTGGGAGAAAGTGTGCTGCTGAAATTTCCAGAAATTTAATTTGATGTTACAT 19217  
QY 2308 AAACCTTTTGACTTCAG 2324  
Db 19218 AAACCTTTTGACTTCAG 19234

## RESULT 8

US-10-109-854-3  
; Sequence 3, Application US/10109854  
; Patent No. 6630337  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL000758DIV  
; CURRENT APPLICATION NUMBER: US/10/109,854  
; CURRENT FILING DATE: 2002-04-01  
; PRIOR APPLICATION NUMBER: 60/227,470  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 09/810,671  
; PRIOR FILING DATE: 2001-03-19

; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 21234  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-109-854-3

Query Match 46.6%; Score 1097; DB 4; Length 21234;  
Best Local Similarity 100.0%; Pred. No. 4.4e-265;  
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1228 AGGAATTTATGCTTTGTGATGATGAAAGCAATGAGAAACTGTTGACCTGGTTCGAAGAA 1287  
Db 18138 AGGAATTTATGCTTTGTGATGATGAAAGCAATGAGAAACTGTTGACCTGGTTCGAAGAA 18197  
QY 1288 TGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGAGAGATCCTTTCT 1347  
Db 18198 TGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGAGAGATCCTTTCT 18257  
QY 1348 TTGACTTTATTAATAAAGAAATGAAATGGGAATCAAGTGGTCTTACTATATATCTTCTCTAGA 1407  
Db 18258 TTGACTTTATTAATAAAGAAATGAAATGGGAATCAAGTGGTCTTACTATATATCTTCTCTAGA 18317  
QY 1408 AGAGATTTACTTAAGACTGTGTCAGTCAACTAAACATTAATCTTAATATTTTGTAAACATTA 1467  
Db 18318 AGAGATTTACTTAAGACTGTGTCAGTCAACTAAACATTAATCTTAATATTTTGTAAACATTA 18377  
QY 1468 TTATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATTAATTAACCTT 1527  
Db 18378 TTATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATTAATTAACCTT 18437  
QY 1528 TTAAGCAAGTATGGCTTGTGATGATGCAATAGAAAAATTAATAATTAATTTCTTTTGGAA 1587  
Db 18438 TTAAGCAAGTATGGCTTGTGATGATGCAATAGAAAAATTAATAATTAATTTCTTTTGGAA 18497  
QY 1588 ATTACCAATTTTAAATACCTTTGAAATATCCCTTGTGTCAGTGATAAATGATGATGATC 1647  
Db 18498 ATTACCAATTTTAAATACCTTTGAAATATCCCTTGTGTCAGTGATAAATGATGATGATC 18557  
QY 1648 TTGCTTTTGTACATGGAGGTCACTCTGAAAGTGAATTTTGTAGTAAAGGAATCTT 1707  
Db 18558 TTGCTTTTGTACATGGAGGTCACTCTGAAAGTGAATTTTGTAGTAAAGGAATCTT 18617  
QY 1708 GACTACTTTATATCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTTA 1767  
Db 18618 GACTACTTTATATCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTTA 18677  
QY 1768 AAAAGTTTTCTTCTGTAATTTCTCAAGGGTGATTTATTTAATCTCTAGATAAGCAGGT 1827  
Db 18678 AAAAGTTTTCTTCTGTAATTTCTCAAGGGTGATTTATTTAATCTCTAGATAAGCAGGT 18737  
QY 1828 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGTGTG 1887  
Db 18738 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGTGTG 18797  
QY 1888 TATTCCTTTTCTATTCGGTGATGTCAGGGTGATAACACAGACATTCATGGAAGGCATGCAG 1947  
Db 18798 TATTCCTTTTCTATTCGGTGATGTCAGGGTGATAACACAGACATTCATGGAAGGCATGCAG 18857  
QY 1948 TTTGTCCATTTGACAGTTTGTAAATTAATAAACACACATACACATTTTAAAGATTAATA 2007  
Db 18858 TTTGTCCATTTGACAGTTTGTAAATTAATAAACACACATACACATTTTAAAGATTAATA 18917  
QY 2008 TCTAACTGGAAGTCACTTGGAAAAATGACATTTTCCAAATGATGTTGGTGAGTCACAGA 2067  
Db 18918 TCTAACTGGAAGTCACTTGGAAAAATGACATTTTCCAAATGATGTTGGTGAGTCACAGA 18977  
QY 2068 TATAAAATAGAAATTTCTGATGAGAGGTTTCCAGTTTTTAAATACCAGTCTTTAGAGTCT 2127  
Db 18978 TATAAAATAGAAATTTCTGATGAGAGGTTTCCAGTTTTTAAATACCAGTCTTTAGAGTCT 19037  
QY 2128 TAACTTTGCGAGCATCTGTTTATCAATGACATAAATACGTAACCTTAAGAATTAAG 2187

```
Db 19038 TAAACATGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTATAAAGAAATTAAG 19097
Qy 2188 TTTATTAAATAGGCAATTTATGCTGATGATGCTTACGGGAGAGAAAGAGGATTTGATTG 2247
Db 19098 TTTATTAAATAGGCAATTTATGCTGATGATGCTTACGGGAGAGAAAGAGGATTTGATTG 19157
Qy 2248 GAAAGCAGTTTGGGAGAAAGTGCTGCTGCAAAATTTCCAGAAATTTAAATTTGATTGGTTACAT 2307
Db 19158 GAAAGCAGTTTGGGAGAAAGTGCTGCTGCAAAATTTCCAGAAATTTAAATTTGATTGGTTACAT 19217
Qy 2308 AAACCTTTTGGACTTCAG 2324
Db 19218 AAACCTTTTGGACTTCAG 19234

RESULT 9
US-10-339-656-3
; Sequence 3, Application US/10339656
; Patent No. 6733978
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000758D1V2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-3

Query Match 46.6%; Score 1097; DB 4; Length 21234;
Best Local Similarity 100.0%; Pred. No. 4.4e-265;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1228 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGACCTGGTTGAGAGAA 1287
Db 18138 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGACCTGGTTGAGAGAA 18197
Qy 1288 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGCGAGCATCCTTTCT 1347
Db 18198 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGCGAGCATCCTTTCT 18257
Qy 1348 TTGACTTTATTAAGAAAGAAATGAAATGGAATCAGTGTCTTACTATATATCTTCTAGA 1407
Db 18258 TTGACTTTATTAAGAAAGAAATGAAATGGAATCAGTGTCTTACTATATATCTTCTAGA 18317
Qy 1408 AGAGATTAATTAAGACTGTGTCAGTCAACTAAACATTTCTAATATATTTTGTAAACATTA 1467
Db 18318 AGAGATTAATTAAGACTGTGTCAGTCAACTAAACATTTCTAATATATTTTGTAAACATTA 18377
Qy 1468 TTATTTTGTACAGTTAAAGTGAATATGATGTTTGTATCAATAGCATATTAATTAACCTG 1527
Db 18378 TTATTTTGTACAGTTAAAGTGAATATGATGTTTGTATCAATAGCATATTAATTAACCTG 18437
Qy 1528 TTAAGCAAGTATGCTTTGATAATGCATTAAGAAATTAAGAAATTAATTTCTTTTTCGAA 1587
Db 18438 TTAAGCAAGTATGCTTTGATAATGCATTAAGAAATTAAGAAATTAATTTCTTTTTCGAA 18497
Qy 1588 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGTGATTGATC 1647
Db 18498 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGTGATTGATC 18557
```

```
RESULT 10
US-09-919-039-238
; Sequence 238, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 238
; LENGTH: 2254
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 420527.25
; NAME/KEY: unsure
```

```
Qy 1648 TTGCTTTTGTACATGGAGGTCACTCTGTAAGTGAATTTTTTTTTCAGTAAAGAAATCTT 1707
Db 18558 TTGCTTTTGTACATGGAGGTCACTCTGTAAGTGAATTTTTTTTTCAGTAAAGAAATCTT 18617
Qy 1708 GACTACTTTATATTTCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAACCTT 1767
Db 18618 GACTACTTTATATTTCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAACCTT 18677
Qy 1768 AAAAGTTTTTCTCTGTAATTTGTTGAACGGGTGATTTATTAATCTAGATAAGCAGT 1827
Db 18678 AAAAGTTTTTCTCTGTAATTTGTTGAACGGGTGATTTATTAATCTAGATAAGCAGT 18737
Qy 1828 ACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTG 1887
Db 18738 ACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTG 18757
Qy 1888 TATTCTTTTTCATTTGGGTGATGTCAGGTGATTAACAGACATTCATGAAAGGCATGAG 1947
Db 18798 TATTCTTTTTCATTTGGGTGATGTCAGGTGATTAACAGACATTCATGAAAGGCATGAG 18857
Qy 1948 TTTCTCCATTTGTGACAGTTTGTAAATAAACCACATACACACTTTATTTAAGATTAAA 2007
Db 18858 TTTCTCCATTTGTGACAGTTTGTAAATAAACCACATACACACTTTATTTAAGATTAAA 18917
Qy 2008 TCTAACTGGAAAGTCAGCTTGAAAAATGGACATTTCCAAAGTATGTTTGGTGAGTCACAGA 2067
Db 18918 TCTAACTGGAAAGTCAGCTTGAAAAATGGACATTTCCAAAGTATGTTTGGTGAGTCACAGA 18977
Qy 2068 TATAAAATAGAAATTCGTATGAGAGTTTCAGTTTTTAAATACCAAGTCTTTAGAGTCT 2127
Db 18978 TATAAAATAGAAATTCGTATGAGAGTTTCAGTTTTTAAATACCAAGTCTTTAGAGTCT 19037
Qy 2128 TAACATTTGGCCAGCATCTGTTTATCAATGACATAAATACCTAATGAATTAAG 2187
Db 19038 TAACATTTGGCCAGCATCTGTTTATCAATGACATAAATACCTAATGAATTAAG 19097
Qy 2188 TTTATTAATTAGGCAATTTATGCTGTGATAAATCTTACGGGAGAAAGAGGATTTGATTG 2247
Db 19098 TTTATTAATTAGGCAATTTATGCTGTGATAAATCTTACGGGAGAAAGAGGATTTGATTG 19157
Qy 2248 GAAAGCAGTTTGGGAGAAAGTGTGCTGAAAATTTCCAGAAATTTAATTTGATTGGTTACAT 2307
Db 19158 GAAAGCAGTTTGGGAGAAAGTGTGCTGAAAATTTCCAGAAATTTAATTTGATTGGTTACAT 19217
Qy 2308 AAACCTTTTGGACTTCAG 2324
Db 19218 AAACCTTTTGGACTTCAG 19234
```

10

Db 623 ACAAGGAAGCGCCGACTAGAAATCAACGTCGGAGAAATCAATGAGAAAGATCTG 682  
Qy 574 ATAGTGTCTTCGGATGTGTCCAGATGCTAGATGTTGATCATCATGCTGATGTTGTA 633  
Db 683 ACAACAGAACTCTGTGTCCAGATGTTGATGCTGTTGATCATCATGCTGATGTTGTA 742  
Qy 634 TTGTTGTTGAACTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAAGAAAGCTTTGCG 693  
Db 743 TCTCTTTGAGCTTCTGGGCTTAGCCTTGCATTTCTCAAGAGCAACAACTACCTGCG 802  
Qy 694 CATTTCAATTTGACCAATCAGGAGATGCGGTATCAGATCTGCCAGTCAATAATTTT 753  
Db 803 CCTACCCCATCACCAAGTGGCCATGCGCTTCCAGCTCTGCCAGGCGCTCAAGTTCC 862  
Qy 754 TACATCAATAAATAAACCCTACAGATCTGAAGCTTGAAGAAATATTTTGTGTAAGT 813  
Db 863 TCCATGATPACAGTTGACATACAGGACCTCAACCTGAAATATTTCTGTTGTGAAT 922  
Qy 814 CTGACTATGTAGTCAATATATTTCTAAATGAAACGTTGATGAACGACATCAAGTACTT 933  
Db 923 CAGACTACGAACTACCTTACCACTTAGAAGAGAGAGATGAGCGCAGTGTAAAGACA 982  
Qy 874 CAGATATCAAGTTGTTGACTTTGAACTTGAAGTGCACCTATGATGATGAACATCACAGTACTT 933  
Db 983 CAGCGTGGGTTGGTGGACTTGGCAGTGGCCATTTTGACCAAGAACACCATAGCACCA 1042  
Qy 934 TGGTCTTACCCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGTGTGCTCAGC 993  
Db 1043 TTGCTCCACTCGCAATTTACCGAGCCCCGAGGTCATCTGGAGTTGGCTGGTCCACAGC 1102  
Qy 994 CTGTGTGTTGGAGCATAGGTTGCATTTCTTATTTGAATATTTACCTGGTTTCAAGTCT 1053  
Db 1103 CATGCGATGTAGGAGCATAGGCTGCATCTTTTGAGTACTAGCTTGGCTTCCACCTCT 1162  
Qy 1054 TTCAGACTCATGATAGTAAGAGCACCTGCAATGATGGAACGATATTTAGGACCCATAC 1113  
Db 1163 TCCAGACCCATGACACAGAGAGCATCTAGCCATGGAAGAGATCTGGGTCCTGTC 1222  
Qy 1114 CACAACATCATGTTTCAAGAAACAGAAACGCAAGTATTTTCAACATAACCAAGCTAGATT 1173  
Db 1223 CTCTCGGATGATCAGAAAGACAGAAACAGAAATATTTTATCGGGTCCCTGGATT 1282  
Qy 1174 GGGATGAACAGATTTCTGCTGGTATGATGTTAGGAGACGCTGCAACCGTTGAAGGAT 1233  
Db 1283 GGGATGAACACCTTCAGCCGCGCTACGTTCTGTGAGAACTGCAAACTCTGCGGCGGT 1342  
Qy 1234 TTATGCTTTGTCATGATCAAGACATGAGAACTGTTTGACCTGTTCCGAAAGATGTTAG 1293  
Db 1343 ATCTGACCTCAGGAGCAGGACCAACCCAGCTCTTCGATCTGATGAAATATGCTAG 1402  
Qy 1294 AATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCTTTTCTTGA 1353  
Db 1403 AGTATGAGCTGCTAAGCGGCTGACCTTAGTGAAGCCCTTCAGCATCTTTCTTGGCT 1462

RESULT 12

US-09-905-999-24  
; Sequence 24, Application US/09905999  
; Patent No. 6797513  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, Axel  
; APPLICANT: NAYLER, Oliver  
; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS  
; FILE REFERENCE: 038602/0431  
; CURRENT APPLICATION NUMBER: US/09/905,999  
; CURRENT FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: 09/127,248  
; PRIOR FILING DATE: 1999-07-31  
; PRIOR APPLICATION NUMBER: PCT/IB97/00946  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 08/877,150  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 60/034,286

; PRIOR FILING DATE: 1996-12-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 24  
; LENGTH: 1787  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-905-999-24  
  
Query Match 19, 3%; Score 454.8; DB 4; Length 1787;  
Best Local Similarity 65.0%; Pred. No. 3.7e-104; Indels 0; Gaps 0;  
Matches 672; Conservative 0; Mismatches 362

Qy 317 CCGAAGGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTGTGATCTGTCAAG 376  
Db 447 CAGTAAGCGCAGCAGCCGAGTGTGAAGATGACAAGGAGGCCACCTGGTGTGCGGAT 506  
Qy 377 TGGAGACGTTCTAAGAGCAAGATATGAATCGTGTGACACTTTGGTGAAGAGCTTTGG 436  
Db 507 CGGCGATTTGGCTCCAAGAGCGATATGATCGTGGGAAACCTGGGTGAAGCACCCTTTGG 566  
Qy 437 CAAAGTTGTAGAGTGCATTTGATCATGATGCGATGCGCATGTAGCAGTGAATAATCGT 496  
Db 567 CAGGTGTGTGAGTGTCTTGGACCATGCCAGAGGGAAGTCAAGGTTGCCCTGAAGATCAT 626  
Qy 497 AAAAAATGTAGCGGTTTACCGTGAAGCAGCTCGTTCAAGAAATCCAAGTATTAGAGCACTT 556  
Db 627 CCGTAATGTGGGCAAGTATCGGGAAGCTGCTCGTCTAGAAATTAATGTCTCAAGAAAT 686  
Qy 557 AAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCCAGATGCTAGATGCTGTTGTATCA 616  
Db 687 CAAAGGAAAGACAAGGAAATAAGTTCTTGTCTCTGATGTCTGACTGGTCAACTT 746  
Qy 617 TCATGTCATCTTTGTTGTTGTTGAACTACTCGGACTTAGTACTTACGATTTCAATAA 676  
Db 747 CCATGTCATATGTGCTGCTTGGCTTTGAGCTCTGGGCAAGAACCTTTGAGTTCTGAA 806  
Qy 677 AGAAAAACAGCTTTCTGCAATTTCAAATTTGACCAATCAGGAGATGCGGTATCAAGATCTG 736  
Db 807 GGAGAACAACTTCCAGCCTTACCCTTACCACATGTCCGGCACATGGCTTACCAGCTCTG 866  
Qy 737 CCAGTCAATAATTTTACATCATATAATAATTAACCATACAGATCTGAAGCCCTGAAA 796  
Db 867 TCATGCCCTTAGATTTCTACACAGAACCCAGCTACCCACACAGACTTGAAGCCAGAGAA 926  
Qy 797 TATTTGTTGTGAAGTCTGACTATGTAGTCAAAATAATTTCTAAAATGAAACCGTATGA 856  
Db 927 CATCTTGTGTGAATTTCTGAGTTTGAACCTCTCAATGAGCACAGAGCTGCGAGGA 986  
Qy 857 ACGCACACTGAAAAACACAGATATCAAAGTTGTTGACTTTGGAAAGTGCACGATATGATGA 916  
Db 987 GAAGTCAGTGAAGAAACACAGCATCCGAGTGGCAGACTTTGGCAGTGCACGTTTGACCA 1046  
Qy 917 TGAACATCAGATGACTTTGGTGTCTACCCGCACTACAGAGCTCCCGAGTCAATTTGGC 976  
Db 1047 TGAACATCAGACCACTTGGCCACCCGTCACCTACCGCCACCTGAGGTGATCTTTGA 1106  
Qy 977 TTTAGGTTGGTCTCAGCCTTGTGATGTTGGAGCATAGTTGCTATCTTATTGAATATTA 1036  
Db 1107 GCTGGCTGGGCACAGCCTTGTGATGCTGGAGTATCGGCTGCATCTCTTTTGAGTACTA 1166  
Qy 1037 CTTTGGTTTCAAGTCTTTTCAAGTCTCATGATAGTAAAGAGACCTTGGCAATGATGGAACG 1096  
Db 1167 CCGTGGCTTTTACACTCTTCCAGACCCATGAAATAGAGAACACTTGGTTATGATGAGGAA 1226  
Qy 1097 AATATTAGGACCCATACCAACACATGATTTCAAGAAACAGAAACCGCAAGTATTTTCA 1156  
Db 1227 GATTCTAGGACCCATCCCATCCACATGATCCACCGTACCAAGGAGCAAGAAATTTTCTA 1286  
Qy 1157 CCATAAACCAAGTATTGGGATGAACAGTCTTCTGCTGGTAGATATGTTAGGAGACGCTG 1216  
Db 1287 CAAAGGGGCTTGGTTTGGGATGAGAACAGCTCTGATGGGCGGTATGTGAAGGAGAACTG 1346

QY 1217 CAACCGTTGAGGAATTTATGCTTTGTCATGATGAAGAAATGAGAACTGTTTGACCT 1276  
DB 1347 CAACCTCTGAAGAGTTACATGCTGACAGATCTCCCTGGAGCATGTCAGCTGTTTGACCT 1406  
QY 1277 GGTTCGAAGAATGTTAGAAATATGATCCAACTCAAGAATTAACCTTGGATGAAGCATTTGCA 1336  
DB 1407 GATGAGGAGGATGTTAGAGTTGACCTTGCACCGCATCACATTTGGCAGAGCCTTGCT 1466  
QY 1337 GCATCTTTCTTTG 1350  
DB 1467 GCACCCCTTCTTTG 1480

## RESULT 13

US-09-016-434-1439  
; Sequence 1439, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1439:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1762 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: 9632971

US-09-016-434-1439  
Query Match 18.4%; Score 432.4; DB 4; Length 1762;  
Best Local Similarity 63.6%; Pred. No. 1.6e-98;  
Matches 658; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 317 CCGAAGGAAAGATCCAGAGTATAGAGATGATGAGAGGGTCACTGATCTGTCAAAG 376  
DB 440 CAGTAAGCGCA CAGCGCGAGTGTGGAAGATGCAAGAGGGTCACTGGTGTGCCGGAT 499  
QY 377 TGGAGACGTTCTAAGAGCAAGATATGAATCGTGACACTTTGGGTGAAGGACCTTTGG 436  
DB 500 CGCGGATTTGGCTCCAGAGCGGATATGAGATTTGTGGGAACTTGGGTGAAGGACCTTTGG 559

QY 437 CARAGTTGTAGAGTGCAATTTGATGCGATGATGGCATGCGATGATGAGTGAATCGT 496  
DB 560 CAAGTGTGTGAGTGTCTTGGACCATGCCAGAGGGAAGTCTCAGGTTGCCCTGAAGATCAT 619  
QY 497 AAAAAATGTAGCGCGTTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGACACTT 556  
DB 620 CCGCAACGTGGCAAGTACCGGAGGCTGCCCGGCTAGAAATCAACGTGCTCAAAAAAT 679  
QY 557 AATAGTACTGATCCCAATAGTGTCTCCGATGTGTCCAGATGCTGAGATGTTGTTGATCA 616  
DB 680 CAAGGAGAAAGCAAAAGAAACAAAGTTCTCTGTGTCTTGTATGCTGCTGCTCAACTT 739  
QY 617 TCATGCTCATGTTTGTATTTGAACTACTCGGACTTGTACTTACGATTTTCAATAA 676  
DB 740 CCACGGTCACATGTGCATCGCTTTGAGCTCTCGGCAAGAACACCTTTTGTAGTTTCTCTGAA 799  
QY 677 AGAAAAACAGCTTTCTGCCATTTCAAAATGACACATCAGGAGATGCGGTATCAGATCTG 736  
DB 800 GGAGATAACTTCCAGCTTACCCCTTACCATGTCCGGCACATGGCTACCGCTCTG 859  
QY 737 CCAGTCAATAAATTTTACATATAATAAATAACCATACAGATCTGAAGCCTGAAAA 796  
DB 860 CCACGCCCTTGAATTTCTGCATGAGAAATCAGCTGACCCATACAGACTTTGAAACCTTGAA 919  
QY 797 TATTTTGTGTGAGTCTGATGTAGTCAATATATATTTCTAAATGAAACCTGTGA 856  
DB 920 CATCTGTTTGTGAATTTCTGAGTTTGAACCCCTCTACAATGAGCACAGAGCTGTGAGGA 979  
QY 857 ACGCACTGAAAAACACAGATATCAAAAGTTGTTGACTTTTGGAACTGCAAGTATGATGA 916  
DB 980 GAAGTCAGTGAAGAACACCAACATCCGAGTGTGCTGACTTTGGCAGTGCACATTTGACCA 1039  
QY 917 TGAACATCACAGTACTTTTGTGTCTTACCCGCACTTACAGAGCTCCCGAGGTCAATTTGGC 976  
DB 1040 TGAGCACCAACACCACTTGTGGCCACCCGCTACTATCGCCGCTGAGGTGATCCTTGA 1099  
QY 977 TTTAGTTGTGCTCAGCTTGTGATGTTTGAGCATAGGTTGCACTTCTTATGATATTA 1036  
DB 1100 GCTGGCTGGGCACAGCCCTGTGAGCTGTGAGCATTTGGCTGCACTTCTTTTGTAGTACTA 1159  
QY 1037 CTTTGGTTTTCACAGTCTTTTACAGACTCATGATAGTAAAGAGCACCTGGCAATGATGAAAG 1096  
DB 1160 CCGGGCTTTCACACTCTTCCAGACCCAGAAACCGAGAGCACCTGGTGTGATGAGGAGNA 1219  
QY 1097 AATATTAGGACCCATACCAACACATGATTTTCAAGAAACAAAGAAACCGCAAGTATTTCOA 1156  
DB 1220 GATCTAGGGCCCATCCCATCACATGATCCCGTACCAGGAGCAGAAATATTTCTA 1279  
QY 1157 CCATAACGAGTATGAGGATGAACACAGTTCTGCTGCTGATATGTTAGGAGCGCTG 1216  
DB 1280 CAAAGGGGGCTAGTTTGGGATGAGAACAGCTCTGACGGCCGCTATGTGAAGGAGAACTG 1339  
QY 1217 CAAACCGTTGAAGGAATTTATGCTTTTGTGATGATGAAGAACATGAGAACTGTTTGACT 1276  
DB 1340 CAAACCTCTGAAGAGTTTACATGCTCCAGAGTCCCTGGAGCACGTGCGAGCTGTTTGACT 1399  
QY 1277 GGTTCGAAGAATGTTAGAAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGA 1336  
DB 1400 GATGAGGAGGATGTTAGAAATTTGACCTGCCAGCGCATCACACTGGCGGAGCCCTGCT 1459  
QY 1337 GCATCTTTCTTTG 1350  
DB 1460 GCACCCCTTCTTTG 1473

## RESULT 14

US-09-949-016-2648  
; Sequence 2648, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF



| Query Match           | 18.38;          | Score 430.8;   | DB 4;     | Length 1763; |
|-----------------------|-----------------|--|-----------|--------------|
| Best Local Similarity | 63.5%;          | Pred. No. 3.9e-98;   |           |              |
| Matches 657;          | Conservative 0; | Mismatches 377;  | Indels 0; | Gaps 0;      |
| Qy                    | 317             | CCGAAGGAAAGATCCAGGAGTAGTAGAGGATGATGAGGAGGTCACCTGATCTCTCAAG     | 376       |              |
| Db                    | 440             | CAGTAACGGCAGACCGCGAGCTGGAGATGACAGGAGGTCACCTGGTGTCCCGAT         | 499       |              |
| Qy                    | 377             | TGGAGACGTTCTTAAGAGCAAGATATGAAATCGTGGACACTTGGGTGAAGAGCGCTTTGG   | 436       |              |
| Db                    | 500             | CGCGATTGGCTCCAAGAGCGATATGAGATTGTGGGGAACCTGGGTGAAGGACACTTTGG    | 559       |              |
| Qy                    | 437             | CAAAAGTTGTAGAGTGCAATTGATCATGCGCATGGATGGCATGCATGTACAGTGAATACTGT | 496       |              |
| Db                    | 560             | CAAGGTGGTGGAGTGTCTTGGACCATGTCAGAGGGAAGTCTCAGGTTCCCTCGAGATCAT   | 619       |              |
| Qy                    | 497             | AAAAATGTAGGCGGTTTACCGTGAAGCAGCTCGTTCAGAAATCCAAGTATTAGAGCACTT   | 556       |              |
| Db                    | 620             | CCGCACGTTGGGCAAGTACCCTGGGAGGTCGCCCGGTAGAAATCAACGTGCTCAAAAAAT   | 679       |              |
| Qy                    | 557             | AAATAGTACTGATCCCAATAGTCTCTCCGATGTGCCAGATGTCAGATGGTTTTGATCA     | 616       |              |
| Db                    | 680             | CAAGGAGAGGACAAAGAAAAACAAGTTCCTGTGTCTTGATGTCGACTGGTTCAACTT      | 739       |              |
| Qy                    | 617             | TCATGGTCATGTTTGATATGTTTGAATCTACTGGGACTTAGTACTTACGATTTCAATAA    | 676       |              |
| Db                    | 740             | CCACGGTCAATGTGATCGCCTTTGAGTCTCTGGGCAAGAACACCTTTGAGTTCCTGAA     | 799       |              |



```
Qy 677 AGAAAAACAGCTTTCTGCCATTTCAAATTTGACCAATCAGGCAGATGGCGTATCAGATCTG 736
Db |||||
Qy 800 GGAGAATAACTTCCAGCCTTACCCCTACCAATGTCGGGCACATGGCTACCAGCTCTG 859
Db |||||
Qy 737 CCAGTCAATAATTTTACATCATATAATAATTAACCCATACAGATCTGAAGCCTGAAAA 796
Db |||||
Qy 860 CCAGCCCTTAGATTTCTGCAATGAGATCAGCTGACCCATACAGACTTTGAAACCAGAGAA 919
Db |||||
Qy 797 TATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAATGAAACGTTGATGA 856
Db |||||
Qy 920 CATCTGTGTGTGATTTCTGAGTTTGAACCTCTACATGAGCACAGAGCTGTGAGGA 979
Db |||||
Qy 857 AGCACACTGAAAAACACAGATATCAAAGTTGTGACTTTTGGAAAGTGCACAGTATGATGA 916
Db |||||
Qy 980 GAAGTCAGTGAAGAACAACAGCATCCGAGTGGCTGACTTTGGCAGTGCACATTTGACCA 1039
Db |||||
Qy 917 TGAACATCACAGTACTTTGGTGTCTACCCGGACCTACAGAGCTCCCGAGGTCAATTTGGC 976
Db |||||
Qy 1040 TGAGCACCAACACCCATTTGTGGCCACCCCTCACTATCCGCCCTGAGGTGATCCTTGA 1099
Db |||||
Qy 977 TTTAGTTGGTCTCAGCCTCTGTGATGTTTGGAGCATAGTTTGCATTTCTATTGAATATTA 1036
Db |||||
Qy 1100 GCTGGCTGGGCACAGCCCTGTGACCTGTGAGCATTTGGCTGCAATCTCTTTGAGTACTA 1159
Db |||||
Qy 1037 CCTTGGTTTCAAGTCTTTTCAGACTCATGATAGTAAGAGCACCTTGGCAATGATGGAACG 1096
Db |||||
Qy 1160 CCGGGCTTTCACACTCTTCCAGACCCACGAAAAACCGAGAGCACCTGGTGTGATGAGAGAA 1219
Db |||||
Qy 1097 AATATTAGGACCCATACACACACATGATTCAGAAACAGAGAAACGCAAGTATTTTCA 1156
Db |||||
Qy 1220 GATCCTAGGGCCCATCCCATCACATGATCCACCGTACCAGGAAGCAGAAATATTTCTA 1279
Db |||||
Qy 1157 CCATAACCAAGTATGATTTGGGATGAACACAGTTCTGCTGTAGATATGTTAGGAGACGCTG 1216
Db |||||
Qy 1280 CAAAGGGGGCTTAGTTTGGGATGAGACAGCTCTGACGGCCGGTATGTGAAGGAGACTG 1339
Db |||||
Qy 1217 CAAACCGTTGAAGGAATTTATGCTTTGTGATGATGAAGAACATGAGAAACTGTTTGACCT 1276
Db |||||
Qy 1340 CAAACCTCTGAAGAGTTTACATGCTCCAAGACTCCCTGGAGCACGTCAGCTGTTTGACCT 1399
Db |||||
Qy 1277 GGTTCGAAGAATGTAGAAATGATGATCCAACTCAAGAATTTACCTTGGATGAAGCATTGCA 1336
Db |||||
Qy 1400 GATGAGGAGGATGTTAGAAATTTGACCCCTGCCCGAGCGCATCACACTGGCCGAGGCCCTGCT 1459
Db |||||
Qy 1337 GCATCCTTTCTTTG 1350
Db |||||
Qy 1460 GCACCCCTTCTTTG 1473
Db |||||
```

Search completed: March 16, 2005, 18:53:25  
Job time : 400.184 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw.model

Run on: March 16, 2005, 14:32:04 ; Search time 1314.36 Seconds  
(without alignments)

10662.144 Million cell updates/sec

Title: US-10-801-671-1

Perfect score: 2354

Sequence: 1 gccagctgggttacttta.....taaaaaaacaaaaaaac 2354

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09D\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description         |
|------------|--------|-------------|--------|-------|---------------------|
| 1          | 2354   | 100.0       | 2354   | 9     | US-09-810-671-1     |
| 2          | 2354   | 100.0       | 2354   | 13    | US-10-109-854-1     |
| 3          | 2354   | 100.0       | 2354   | 15    | US-10-339-656-1     |
| 4          | 2354   | 100.0       | 2354   | 18    | US-10-801-671-1     |
| 5          | 1908.4 | 81.1        | 4035   | 17    | US-10-425-114-26852 |
| 6          | 1286   | 54.6        | 1446   | 17    | US-10-267-502-135   |
| 7          | 1178.2 | 50.1        | 1456   | 17    | US-10-641-643-699   |
| 8          | 1097.4 | 46.6        | 1549   | 9     | US-09-905-999-26    |
| 9          | 1097.4 | 46.6        | 1549   | 18    | US-10-825-177-26    |
| 10         | 1097.2 | 46.6        | 1446   | 17    | US-10-267-502-138   |
| 11         | 1097   | 46.6        | 21234  | 9     | US-09-810-671-3     |

|    |       |      |       |    |                     |                   |
|----|-------|------|-------|----|---------------------|-------------------|
| 12 | 1097  | 46.6 | 21234 | 13 | US-10-109-854-3     | Sequence 3, Appli |
| 13 | 1097  | 46.6 | 21234 | 15 | US-10-339-656-3     | Sequence 3, Appli |
| 14 | 1097  | 46.6 | 21234 | 18 | US-10-801-671-3     | Sequence 3, Appli |
| 15 | 977.4 | 41.5 | 3040  | 17 | US-10-425-114-26212 | Sequence 26212, A |
| 16 | 813.2 | 34.5 | 1455  | 17 | US-10-267-502-136   | Sequence 136, App |
| 17 | 811.6 | 34.5 | 1834  | 17 | US-10-154-708-3     | Sequence 3, Appli |
| 18 | 811.6 | 34.5 | 1834  | 18 | US-10-755-889-1     | Sequence 1, Appli |
| 19 | 768.8 | 32.7 | 1452  | 17 | US-10-267-502-139   | Sequence 139, App |
| 20 | 768.4 | 32.6 | 906   | 17 | US-10-267-502-140   | Sequence 140, App |
| 21 | 644.4 | 27.4 | 1743  | 17 | US-10-154-708-10    | Sequence 10, Appl |
| 22 | 640.6 | 27.2 | 2254  | 10 | US-09-919-039-238   | Sequence 238, App |
| 23 | 640   | 27.2 | 2516  | 9  | US-09-925-298-121   | Sequence 121, App |
| 24 | 640   | 27.2 | 2516  | 14 | US-10-102-806-121   | Sequence 121, App |
| 25 | 499.6 | 21.2 | 1500  | 10 | US-09-790-852-2     | Sequence 2, Appli |
| 26 | 499.6 | 21.2 | 1973  | 9  | US-09-962-436-266   | Sequence 266, App |
| 27 | 499.6 | 21.2 | 1973  | 9  | US-09-880-107-2190  | Sequence 2190, Ap |
| 28 | 499.6 | 21.2 | 1973  | 18 | US-10-737-450-131   | Sequence 131, App |
| 29 | 481.2 | 20.4 | 1538  | 9  | US-09-905-999-22    | Sequence 22, Appl |
| 30 | 481.2 | 20.4 | 1538  | 18 | US-10-825-177-22    | Sequence 22, Appl |
| 31 | 464.4 | 19.7 | 1296  | 17 | US-10-182-243-16    | Sequence 16, Appl |
| 32 | 464.4 | 19.7 | 1628  | 17 | US-10-425-114-26266 | Sequence 26266, A |
| 33 | 461.6 | 19.6 | 1026  | 17 | US-10-267-502-133   | Sequence 133, App |
| 34 | 458.6 | 19.5 | 1885  | 17 | US-10-439-703-89    | Sequence 89, Appl |
| 35 | 454.8 | 19.3 | 1787  | 9  | US-09-905-999-24    | Sequence 24, Appl |
| 36 | 454.8 | 19.3 | 1787  | 18 | US-10-825-177-24    | Sequence 24, Appl |
| 37 | 453.2 | 19.3 | 1473  | 17 | US-10-267-502-137   | Sequence 137, App |
| 38 | 432.4 | 18.4 | 1473  | 17 | US-10-267-502-134   | Sequence 134, App |
| 39 | 432.4 | 18.4 | 1762  | 17 | US-10-305-720-1439  | Sequence 1439, Ap |
| 40 | 430.8 | 18.3 | 2505  | 17 | US-10-108-260A-2256 | Sequence 2256, Ap |
| 41 | 429.2 | 18.2 | 2621  | 17 | US-10-104-047-656   | Sequence 656, App |
| 42 | 377.4 | 16.0 | 1181  | 17 | US-10-425-114-16170 | Sequence 16170, A |
| 43 | 369   | 15.7 | 369   | 9  | US-09-796-692-5456  | Sequence 5456, Ap |
| 44 | 369   | 15.7 | 369   | 14 | US-10-040-862-5456  | Sequence 5456, Ap |
| 45 | 369   | 15.7 | 369   | 17 | US-10-057-475B-5456 | Sequence 5456, Ap |

ALIGNMENTS

RESULT 1  
US-09-810-671-1  
; Sequence 1, Application US/09810671  
; Publication No. US20020076783A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CLO00758  
; CURRENT APPLICATION NUMBER: US/09/810.671  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2354  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-810-671-1

|                       |         |                      |  |            |    |        |       |
|-----------------------|---------|----------------------|--|------------|----|--------|-------|
| Query Match           | 100.0%; | Score                | 2354;                                    | DB         | 9; | Length | 2354; |
| Best Local Similarity | 100.0%; | Pred. No.            | 0;                                       |            |    |        |       |
| Matches               | 2354;   | Conservative         | 0;                                       | Mismatches | 0; | Indels | 0;    |
|                       |         |                      |  |            |    |        |       |
| Qy                    | 1       | GCCAGCTGGGTACTTTTAA  | AAACATGCTCCATGTGATCCCTCTTGAAGCTTCGCACT   | 60         |    |        |       |
| Dd                    | 1       | GCCAGCTGGGTACTTTTAA  | AAACATGCTCCATGTGATCCCTCTTGAAGCTTCGCACT   | 60         |    |        |       |
| Qy                    | 61      | CTGTTGAAGAGGACACTCAT | CCCCAGTCAATTTTAGAAGCAAGGTCCTTGAATGAGCGAG | 120        |    |        |       |
| Dd                    | 61      | CTGTTGAAGAGGACACTCAT | CCCCAGTCAATTTTAGAAGCAAGGTCCTTGAATGAGCGAG | 120        |    |        |       |
| Qy                    | 121     | ATTATCGGACCGGAGATAC  | GTTCAGTAATCAGGAATGACTACTGTGAGGATATGTC    | 180        |    |        |       |
| Dd                    | 121     | ATTATCGGACCGGAGATAC  | GTTCAGTAATCAGGAATGACTACTGTGAGGATATGTC    | 180        |    |        |       |

|     |   |      |
|-----|---|------|
| 121 | ATTATCGGGACCGGAGATACGTTGACGAAATACAGAAATGACTACTGTGGAAGGATATGTTCT | 180  |
| Qy  | CTAGACATTATCACAGAGACATTCGAAAGCGGGTATCGAATCCACTGACGTAAATCTTCACG  | 240  |
|     | CTAGACATTATCACAGAGACATTCGAAAGCGGGTATCGAATCCACTGACGTAAATCTTCACG  | 240  |
|     | CTAGACATTATCACAGAGACATTCGAAAGCGGGTATCGAATCCACTGACGTAAATCTTCACG  | 240  |
| Qy  | TCCCGACGAGGAGACAGTCCTTAAAGGAAGCGCAATAGACACTGTGTTCAAGTCATCAGT    | 300  |
| Db  | TCCCGACGAGGAGAGCAGTCCTTAAAGGAAGCGCAATAGACACTGTGTTCAAGTCATCAGT   | 300  |
| Qy  | CACGTTGMAAGACCGACCGAGGAAGAAAGTCCACGAGCTATAGAGGATGATGAGAGAGGTC   | 360  |
| Db  | CACGTTGMAAGACCGACCGAGGAAGAAAGTCCACGAGCTATAGAGGATGATGAGAGAGGTC   | 360  |
| Qy  | ACCTGATCTGTCAAAGTGGAGACGCTTCTTAAGAGCAAGATATGAAATCGTGGACACTTTGG  | 420  |
| Db  | ACCTGATCTGTCAAAGTGGAGACGCTTCTTAAGAGCAAGATATGAAATCGTGGACACTTTGG  | 420  |
| Qy  | GTGAAGGAGCCTTTTGGCAAAAGTTGTAGAGTGCAATGATCATGGCATGGATGGCATGTCATG | 480  |
| Db  | GTGAAGGAGCCTTTTGGCAAAAGTTGTAGAGTGCAATGATCATGGCATGGATGGCATGTCATG | 480  |
| Qy  | TAGCAGTGAATAATCGTAAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCC   | 540  |
| Db  | TAGCAGTGAATAATCGTAAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCC   | 540  |
| Qy  | AAGTATTAGAGCACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCCAGATGC     | 600  |
| Db  | AAGTATTAGAGCACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCCAGATGC     | 600  |
| Qy  | TAGAAATGGTTTGATCATCATGGTTCATGTTGTATTGTGTTTGAACACTACTGGGACTTAGTA | 660  |
| Db  | TAGAAATGGTTTGATCATCATGGTTCATGTTGTATTGTGTTTGAACACTACTGGGACTTAGTA | 660  |
| Qy  | CTTAGCATTTCAATTAAGAAAAAGCTTTTCTGCCATTTCAAATGACCAATACAGGCAGA     | 720  |
| Db  | CTTAGCATTTCAATTAAGAAAAAGCTTTTCTGCCATTTCAAATGACCAATACAGGCAGA     | 720  |
| Qy  | TGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAG     | 780  |
| Db  | TGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAG     | 780  |
| Qy  | ATCTGAAGCCTGAAAAATATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTA   | 840  |
| Db  | ATCTGAAGCCTGAAAAATATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTA   | 840  |
| Qy  | AAATGAACCGTGATGAAGCCACACTGAAAAACACAGATATCAAAGTTGTTGACCTTTGGA    | 900  |
| Db  | AAATGAACCGTGATGAAGCCACACTGAAAAACACAGATATCAAAGTTGTTGACCTTTGGA    | 900  |
| Qy  | GTGCAACGTATGATGATGAACATCACAGTACTTTGGTGTCTACCCGACACTACAGAGCTC    | 960  |
| Db  | GTGCAACGTATGATGATGAACATCACAGTACTTTGGTGTCTACCCGACACTACAGAGCTC    | 960  |
| Qy  | CCGAGGTCATTTTGGCTTTAGTTTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGCA   | 1020 |
| Db  | CCGAGGTCATTTTGGCTTTAGTTTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGCA   | 1020 |
| Qy  | TTCTTATTAAGATATTAACCTTGGTTTCCAGTCTTTTCCAGTCTCATGATAGTAAGAGCACC  | 1080 |
| Db  | TTCTTATTAAGATATTAACCTTGGTTTCCAGTCTTTTCCAGTCTCATGATAGTAAGAGCACC  | 1080 |
| Qy  | TGGCAATGATGGGAACGAATATTAGGACCCATACCACAAACATGATTCAGAAACAGAA      | 1140 |
| Db  | TGGCAATGATGGGAACGAATATTAGGACCCATACCACAAACATGATTCAGAAACAGAA      | 1140 |
| Qy  | AACGCAAGTATTTTACCATAAACAGCTAGATTTGGGATGAACACAGTTCCTGCTGGTAGAT   | 1200 |
| Db  | AACGCAAGTATTTTACCATAAACAGCTAGATTTGGGATGAACACAGTTCCTGCTGGTAGAT   | 1200 |
| Qy  | ATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAAACATG  | 1260 |
| Db  | ATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAAACATG  | 1260 |

|    |      |  |      |
|----|------|--|------|
| QY | 1261 | AGAAACTGTTTGACCTGGTTCGAAGAAATGTTAGAAATATGATCOACATCAAGAAATTTACCT  | 1322 |
| DB | 1261 | AGAAACTGTTTGACCTGGTTCGAAGAAATGTTAGAAATATGATCOACATCAAGAAATTTACCT  | 1320 |
| QY | 1321 | TGGATGAAGCATTTGCAGCATCCCTTCTTTCGACATTATTTAAAAAGAAATGAAATGCGGAATC | 1380 |
| DB | 1321 | TGGATGAAGCATTTGCAGCATCCCTTCTTTCGACATTATTTAAAAAGAAATGAAATGCGGAATC | 1380 |
| QY | 1381 | AGTGGTCTTACTATATATCTTCTCTAGAAGAGATTACTTTAAGACATGTGTCACTCAACTAAA  | 1440 |
| DB | 1381 | AGTGGTCTTACTATATATCTTCTCTAGAAGAGATTACTTTAAGACATGTGTCACTCAACTAAA  | 1440 |
| QY | 1441 | CATTCTTAATATTTTGTAAACATTAATATTTTGTACAGTTAAAGTGAATAATTTCTGATG     | 1500 |
| DB | 1441 | CATTCTTAATATTTTGTAAACATTAATATTTTGTACAGTTAAAGTGAATAATTTCTGATG     | 1500 |
| QY | 1501 | TTTTTGATCAATAGCATAAATTAACCTTGTTTAAGCAAGTATGCTCTTGATATGCAATTAGAA  | 1560 |
| DB | 1501 | TTTTTGATCAATAGCATAAATTAACCTTGTTTAAGCAAGTATGCTCTTGATATGCAATTAGAA  | 1560 |
| QY | 1561 | AAATTAATAATTTTCTTTTTTGAAATTTACCATTTTAAATACCTTTGAAATATCCCTT       | 1620 |
| DB | 1561 | AAATTAATAATTTTCTTTTTTGAAATTTACCATTTTAAATACCTTTGAAATATCCCTT       | 1620 |
| QY | 1621 | TGCTGCAGTGATAAATGTCATTTGATCTTGCCCTTTTGTACATCGAGGTCACTCTCAAGT     | 1680 |
| DB | 1621 | TGCTGCAGTGATAAATGTCATTTGATCTTGCCCTTTTGTACATCGAGGTCACTCTCAAGT     | 1680 |
| QY | 1681 | GATTTTTTTTTGAGTAAAGGAAATCTTGACTACTTTATATCTTTAAAGGAATATCTTTTA     | 1740 |
| DB | 1681 | GATTTTTTTTTGAGTAAAGGAAATCTTGACTACTTTATATCTTTAAAGGAATATCTTTTA     | 1740 |
| QY | 1741 | TATACTTTCAAAATTTAGAACTTTAAAGTTTTTCTCTGTAATCTGTGAACGGGTG          | 1800 |
| DB | 1741 | TATACTTTCAAAATTTAGAACTTTAAAGTTTTTCTCTGTAATCTGTGAACGGGTG          | 1800 |
| QY | 1801 | ATTATTTAACTCTAGATAAGCAGGTACTAGAAAAACCAAACTCAGAAAAATGTTTACTGT     | 1860 |
| DB | 1801 | ATTATTTAACTCTAGATAAGCAGGTACTAGAAAAACCAAACTCAGAAAAATGTTTACTGT     | 1860 |
| QY | 1861 | TAGAAATCTATTAATAATTTTAAAGTTGTATCTTTTTTTCATTTGGGTGATGTCAGGGTGATA  | 1920 |
| DB | 1861 | TAGAAATCTATTAATAATTTTAAAGTTGTATCTTTTTTTCATTTGGGTGATGTCAGGGTGATA  | 1920 |
| QY | 1921 | ACCAGACATTCATGGAAGGCATGCAAGTTGTCCATTTGTGACAGTTTGTTTTAAATAAAACC   | 1980 |
| DB | 1921 | ACCAGACATTCATGGAAGGCATGCAAGTTGTCCATTTGTGACAGTTTGTTTTAAATAAAACC   | 1980 |
| QY | 1981 | ACATACACATTTTATTTAAGATTAAATCTAACTGGAAAGTCAAGTTGGAAAAATGGACAT     | 2040 |
| DB | 1981 | ACATACACATTTTATTTAAGATTAAATCTAACTGGAAAGTCAAGTTGGAAAAATGGACAT     | 2040 |
| QY | 2041 | TTCCAGATGTTTGGTGAGTCAAGATATAAAATAGAAAAATCTCATGACAGAGTTTCAG       | 2100 |
| DB | 2041 | TTCCAGATGTTTGGTGAGTCAAGATATAAAATAGAAAAATCTCATGACAGAGTTTCAG       | 2100 |
| QY | 2101 | TTTTTTAATACCAAGTCTTTAGGAGTCTTAACTATGGCCAGCATCTGTTTATCAAAATGACA   | 2160 |
| DB | 2101 | TTTTTTAATACCAAGTCTTTAGGAGTCTTAACTATGGCCAGCATCTGTTTATCAAAATGACA   | 2160 |
| QY | 2161 | TAAATACGTAAACCTATAAGAAATTAAGTTTAAATTTAGGCAATTTATGTCGTGATAAT      | 2220 |
| DB | 2161 | TAAATACGTAAACCTATAAGAAATTAAGTTTAAATTTAGGCAATTTATGTCGTGATAAT      | 2220 |
| QY | 2221 | TCCTACGGGAGAAAGAGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAAT     | 2280 |
| DB | 2221 | TCCTACGGGAGAAAGAGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAAT     | 2280 |
| QY | 2281 | TTCCAGAAATTTAATGATTTGTTTATATAAACTTTTGTGACTTCAGAAAAAAAATAATAAAA   | 2340 |
| DB | 2281 | TTCCAGAAATTTAATGATTTGTTTATATAAACTTTTGTGACTTCAGAAAAAAAATAATAAAA   | 2340 |

QY 2341 AACAAAAAAMAC 2354  
DB 2341 AACAAAAAAMAC 2354

## RESULT 2

US-10-109-854-1  
; Sequence 1, Application US/10109854  
; Publication No. US20020119548A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THERMOF  
; FILE REFERENCE: CLO00758DIV  
; CURRENT APPLICATION NUMBER: US/10/109,854  
; CURRENT FILING DATE: 2002-04-01  
; PRIOR APPLICATION NUMBER: 60/227,470  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 09/810,671  
; PRIOR FILING DATE: 2001-03-19  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PaetSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2354  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-109-854-1

Query Match 100.0%; Score 2354; DB 13; Length 2354;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGCTGGGTTACTTTAAAAAACATGCTCCATGTCATCCCTCTCGAAGCTTCGCACT 60  
DB 1 GCCAGCTGGGTTACTTTAAAAAACATGCTCCATGTCATCCCTCTCGAAGCTTCGCACT 60

QY 61 CTGTTGAAGAGGACACTCATCCAGTCATTTAAGAGCAAGGTCCTTGAATGAGCGAG 120  
DB 61 CTGTTGAAGAGGACACTCATCCAGTCATTTAAGAGCAAGGTCCTTGAATGAGCGAG 120

QY 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTGTCGAGGATATGTC 180  
DB 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTGTCGAGGATATGTC 180

QY 181 CTAGACATTTATCACAGAGACATTTGAAAGCGGTATCGAATCCACTGCAATTAATCTTCAG 240  
DB 181 CTAGACATTTATCACAGAGACATTTGAAAGCGGTATCGAATCCACTGCAATTAATCTTCAG 240

QY 241 TCCGACGAGGAGAGAGCTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGT 300  
DB 241 TCCGACGAGGAGAGAGCTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGT 300

QY 301 CACGTTCAAGAGCCACCGAAGAAAGATCAGAGATATAGAGGATGATGAGGAGGTC 360  
DB 301 CACGTTCAAGAGCCACCGAAGAAAGATCAGAGATATAGAGGATGATGAGGAGGTC 360

QY 361 ACCTGATCTGTCAAAGTGGAGACGTTCTAAGAGCAAGATATGAAATCGTGGACACTTTGG 420  
DB 361 ACCTGATCTGTCAAAGTGGAGACGTTCTAAGAGCAAGATATGAAATCGTGGACACTTTGG 420

QY 421 GTGAAGGAGCCCTTTGGCAAGGTTGTAGAGTGCAATTCATGATCGCATGATGATGATG 480  
DB 421 GTGAAGGAGCCCTTTGGCAAGGTTGTAGAGTGCAATTCATGATCGCATGATGATGATG 480

QY 481 TAGCAGTGAATCTGTAATAAATGTAGCGGTTACCGTGAAGCAGCTGTTCCAGAAATCC 540  
DB 481 TAGCAGTGAATCTGTAATAAATGTAGCGGTTACCGTGAAGCAGCTGTTCCAGAAATCC 540

QY 541 AAGTATTAGACACATTAATAGTACTGATCCCAATAGTGTCTCCGATGTCGACAGATGC 600  
DB 541 AAGTATTAGACACATTAATAGTACTGATCCCAATAGTGTCTCCGATGTCGACAGATGC 600

QY 601 TAGAATGGTTTGATCATCATGGTCAATGTTGTGTAATTTGTTGAACTACTGGGACTTAGTA 660  
DB 601 TAGAATGGTTTGATCATCATGGTCAATGTTGTGTAATTTGTTGAACTACTGGGACTTAGTA 660

QY 661 CTTACGATTTTCATTTAAAGAAACAGCTTTCTGCCAATTTCAAAATTCACCAATCAGGAG 720  
DB 661 CTTACGATTTTCATTTAAAGAAACAGCTTTCTGCCAATTTCAAAATTCACCAATCAGGAG 720

QY 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTATCATATAATAAATTAACCCATACAG 780  
DB 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTATCATATAATAAATTAACCCATACAG 780

QY 781 ATCTGAAGCCTGAAATAATTTTGTGTAAGTCTGACTATGTAGTCAAAATATAATTTCTA 840  
DB 781 ATCTGAAGCCTGAAATAATTTTGTGTAAGTCTGACTATGTAGTCAAAATATAATTTCTA 840

QY 841 AAATGAACGCTGATGAACGACACTGAAACACACAGATATCAAAAGTTCTGACCTTTGAA 900  
DB 841 AAATGAACGCTGATGAACGACACTGAAACACACAGATATCAAAAGTTCTGACCTTTGAA 900

QY 901 GTGCAACGATATGATGATGAACATCACAGTACTTTTGGTGTCTACCCGGCACTACAGAGTC 960  
DB 901 GTGCAACGATATGATGATGAACATCACAGTACTTTTGGTGTCTACCCGGCACTACAGAGTC 960

QY 961 CCGAGGCTATTTTGGCTTTAGTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGCA 1020  
DB 961 CCGAGGCTATTTTGGCTTTAGTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGCA 1020

QY 1021 TTTCTATTGAATATTTACCTTGGTTTACAGTCTTTTACAGTCTCATGATAGTAAAGACACC 1080  
DB 1021 TTTCTATTGAATATTTACCTTGGTTTACAGTCTTTTACAGTCTCATGATAGTAAAGACACC 1080

QY 1081 TGGCAATGATGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAAACAAGAA 1140  
DB 1081 TGGCAATGATGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAAACAAGAA 1140

QY 1141 AACGCAAGTATTTTCAACATTAACAGCTAGATTTGGATGAACACAGTTCGCTGGTAGAT 1200  
DB 1141 AACGCAAGTATTTTCAACATTAACAGCTAGATTTGGATGAACACAGTTCGCTGGTAGAT 1200

QY 1201 ATGTTAGGAGACGCTGCAACCGTTGAGGAATTTATGCTTTGTCATCATGAAGAACATG 1260  
DB 1201 ATGTTAGGAGACGCTGCAACCGTTGAGGAATTTATGCTTTGTCATCATGAAGAACATG 1260

QY 1261 AGAACTGTTTGAACCTGGTTCGAAGAATTTAGAAATATGATCCAACTCAAGAAATACCT 1320  
DB 1261 AGAACTGTTTGAACCTGGTTCGAAGAATTTAGAAATATGATCCAACTCAAGAAATACCT 1320

QY 1321 TGGATGAAGCATTTGAGCATCCTTTCTTGTGATTTAATAAAGAAATGAAATGGGAATC 1380  
DB 1321 TGGATGAAGCATTTGAGCATCCTTTCTTGTGATTTAATAAAGAAATGAAATGGGAATC 1380

QY 1381 AGTGGTCTTACTATATATCTCTTAGAGAGATTTTAAAGCTGTGTCACTCAACTAAA 1440  
DB 1381 AGTGGTCTTACTATATATCTCTTAGAGAGATTTTAAAGCTGTGTCACTCAACTAAA 1440

QY 1441 CATTTCTAATATTTTGTAAACATTTAAATTTTGTACAGTTAAGTGTAAATTTGTATG 1500  
DB 1441 CATTTCTAATATTTTGTAAACATTTAAATTTTGTACAGTTAAGTGTAAATTTGTATG 1500

QY 1501 TTTTGTATCAATAGCATTAATACTTTGTAAAGCAAGTATGCTTTGTGATTAATGATTAAG 1560  
DB 1501 TTTTGTATCAATAGCATTAATACTTTGTAAAGCAAGTATGCTTTGTGATTAATGATTAAG 1560

QY 1561 AAATTTAAATTTTCTTTTGAATTAACATTTTAAATACCTTTTGAATATCCCTT 1620  
DB 1561 AAATTTAAATTTTCTTTTGAATTTAAATTTTAAATACCTTTTGAATATCCCTT 1620

QY 1621 TGTGTCAGATGATAAATGATGATCTTGTGCTTTTGTACATGAGGAGTCACTCTGGAAGT 1680  
DB 1621 TGTGTCAGATGATAAATGATGATCTTGTGCTTTTGTACATGAGGAGTCACTCTGGAAGT 1680

QY 1681 GATTTTGTGAGTAAAGGAAATCTTGCATCTTATATCTTTAAAGGAAATATTTCTTTA 1740

|    |      |  |      |
|----|------|--|------|
| Db | 1581 | GATTTTTTTGGTAAAGGAAATCTTGACATCTTATATCTTTAAAGGAATATCTCTTA         | 1740 |
| Qy | 1741 | TATACTTCAAAATTTAGAACTTAACTTTAAAGATTTTTCTCTGTAAATGTTGAACGGGTG     | 1800 |
| Db | 1741 | TATACTTCAAAATTTAGAACTTAACTTTAAAGATTTTTCTCTGTAAATGTTGAACGGGTG     | 1800 |
| Qy | 1801 | ATTATATTAACTCTAGATAGCAGGTACTAGAAACCAAAACTCAGAAAATGTTTACTGT       | 1860 |
| Db | 1801 | ATTATATTAACTCTAGATAGCAGGTACTAGAAACCAAAACTCAGAAAATGTTTACTGT       | 1860 |
| Qy | 1861 | TAGAAATCTATTAATTTAAGTGTTGTATTTCTTTTTCAITGGGTGATGTCAGGGTGATA      | 1920 |
| Db | 1861 | TAGAAATCTATTAATTTAAGTGTTGTATTTCTTTTTCAITGGGTGATGTCAGGGTGATA      | 1920 |
| Qy | 1921 | ACCAGACATTCATGGAAGGCATGCAGTTTGTCCATTTGCACAGTTTGTTTAATAAAACC      | 1980 |
| Db | 1921 | ACCAGACATTCATGGAAGGCATGCAGTTTGTCCATTTGCACAGTTTGTTTAATAAAACC      | 1980 |
| Qy | 1981 | ACATACACACTTTATTTAAGATTAAAAATCAAACCTGGAAAGTCAGCTTTGGAAAAATGGACAT | 2040 |
| Db | 1981 | ACATACACACTTTATTTAAGATTAAAAATCAAACCTGGAAAGTCAGCTTTGGAAAAATGGACAT | 2040 |
| Qy | 2041 | TTCCAAAGTGTGTGGTGAGTCACAGATATAAAAATAGAAATTTCTGATGAGAGGTTTCAG     | 2100 |
| Db | 2041 | TTCCAAAGTGTGTGGTGAGTCACAGATATAAAAATAGAAATTTCTGATGAGAGGTTTCAG     | 2100 |
| Qy | 2101 | TTTTTAAACCAAGTCCTTTAGGAGTCCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACA  | 2160 |
| Db | 2101 | TTTTTAAACCAAGTCCTTTAGGAGTCCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACA  | 2160 |
| Qy | 2161 | TAAATAGTAAACCTATAAGAAATTAAGTTTATTAATTTAGGCAATTTATGTCGTGATAAT     | 2220 |
| Db | 2161 | TAAATAGTAAACCTATAAGAAATTAAGTTTATTAATTTAGGCAATTTATGTCGTGATAAT     | 2220 |
| Qy | 2221 | TCTTTACGGGAGAAAGAGGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTCCTCTCAAAAT    | 2280 |
| Db | 2221 | TCTTTACGGGAGAAAGAGGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTCCTCTCAAAAT    | 2280 |
| Qy | 2281 | TTCCAGAAATTTAATGATGGTTTACATAAACTTTTTTGACTTTCAGAAAAAAAATAAAAA     | 2340 |
| Db | 2281 | TTCCAGAAATTTAATGATGGTTTACATAAACTTTTTTGACTTTCAGAAAAAAAATAAAAA     | 2340 |
| Qy | 2341 | AACAAAAAATAAAC 2354  |      |
| Db | 2341 | AACAAAAAATAAAC 2354  |      |

### RESULT 3

```

US-10-339-656-1
; Sequence 1, Application US/10339656
; Publication No. US20030134319A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-1

```

QY 1021 TTCCTATTGAATATTTACCTTGGTTTTCAGAGTCTTTCAGAGTCTATGATAGTAAGAGCACC 1080  
 Db 1021 TTCCTATTGAATATTTACCTTGGTTTTCAGAGTCTTTCAGAGTCTATGATAGTAAGAGCACC 1080  
 QY 1081 TGGCAATGATGGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAACAGAA 1140  
 Db 1081 TGGCAATGATGGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAACAGAA 1140  
 QY 1141 AACGCAAGTATTTTCCACATTAACAGCTAGATTTGGATGAACACAGCTTCTGCTGGTAGAT 1200  
 Db 1141 AACGCAAGTATTTTCCACATTAACAGCTAGATTTGGATGAACACAGCTTCTGCTGGTAGAT 1200  
 QY 1201 ATGTTAGAGAGCGCTGCAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATG 1260  
 Db 1201 ATGTTAGAGAGCGCTGCAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATG 1260  
 QY 1261 AGAACTGTTTGAACCTGTTGCAAGAAATGTTAGAAATATGATCCAACTCAAAAGAAATTAACCT 1320  
 Db 1261 AGAACTGTTTGAACCTGTTGCAAGAAATGTTAGAAATATGATCCAACTCAAAAGAAATTAACCT 1320  
 QY 1321 TGGATGAAGCATTCAGCATCTTCTTTTGACTTATTAATAAAGAAATGAAATGGGAATC 1380  
 Db 1321 TGGATGAAGCATTCAGCATCTTCTTTTGACTTATTAATAAAGAAATGAAATGGGAATC 1380  
 QY 1381 AGTGTCTTACTATATCTTCTAGAGAGATTTAAAGCTGTGTCAGTCAACTTAA 1440  
 Db 1381 AGTGTCTTACTATATCTTCTAGAGAGATTTAAAGCTGTGTCAGTCAACTTAA 1440  
 QY 1441 CATCTCAATATTTTGTAAACATTAATATTTTGTACAGTTAAAGTGAATATTTGATG 1500  
 Db 1441 CATCTCAATATTTTGTAAACATTAATATTTTGTACAGTTAAAGTGAATATTTGATG 1500  
 QY 1501 TTTTGTATCAATAGCATTAATTAACCTTGTAAAGCAAGTATGCTTTGATAATGCAATGAA 1560  
 Db 1501 TTTTGTATCAATAGCATTAATTAACCTTGTAAAGCAAGTATGCTTTGATAATGCAATGAA 1560  
 QY 1561 AAATTTAAATTTTCTTTTGTAAATTAACCTTTTAAATACCTTTGAAATATCCTT 1620  
 Db 1561 AAATTTAAATTTTCTTTTGTAAATTTAAATTTAAATTTAAATTTGAAATATCCTT 1620  
 QY 1621 TGTGTCAGTATGATTAATCTGATGATCTTGTGCTTTTGTACATGAGGTCACCTCTGAAGT 1680  
 Db 1621 TGTGTCAGTATGATTAATCTGATGATCTTGTGCTTTTGTACATGAGGTCACCTCTGAAGT 1680  
 QY 1681 GATTTTTTTTGAAGTAAAGGAAATCTTGACTACTTTTATTTTCTTAAAGCAATATCTTTA 1740  
 Db 1681 GATTTTTTTTGAAGTAAAGGAAATCTTGACTACTTTTATTTTCTTAAAGCAATATCTTTA 1740  
 QY 1741 TATACCTCAAAATTTAGAACTTAACTTTTAAAGTCTTCTGTAATTTGTTGAACGGGTG 1800  
 Db 1741 TATACCTCAAAATTTAGAACTTAACTTTTAAAGTCTTCTGTAATTTGTTGAACGGGTG 1800  
 QY 1801 ATTATTTTAACTCTAGATAGCAGGTACTAGAACCAAACTCAGAAAATGTTTACTGT 1860  
 Db 1801 ATTATTTTAACTCTAGATAGCAGGTACTAGAACCAAACTCAGAAAATGTTTACTGT 1860  
 QY 1861 TAGAATCTTATTAATTTTAAAGTGTGTTATCTTTTCAATTTGGGTGATGTCAGGCTGATA 1920  
 Db 1861 TAGAATCTTATTAATTTTAAAGTGTGTTATCTTTTCAATTTGGGTGATGTCAGGCTGATA 1920  
 QY 1921 ACCAGACATTCATGAAAGGCAATGAGTTTGTGCAATTTGTCAGAGTTTGTAAATAAAC 1980  
 Db 1921 ACCAGACATTCATGAAAGGCAATGAGTTTGTGCAATTTGTCAGAGTTTGTAAATAAAC 1980  
 QY 1981 ACATACACATTTTAAAGTATTAATCTTAACCTGGAAGTCAGCTTGGAAAATGGAAT 2040  
 Db 1981 ACATACACATTTTAAAGTATTAATCTTAACCTGGAAGTCAGCTTGGAAAATGGAAT 2040  
 QY 2041 TTCCAAGTATGTTGGTGAAGTACAGATATAAATAAGAAATTTCTGATGAGAGGTTTCAG 2100  
 Db 2041 TTCCAAGTATGTTGGTGAAGTACAGATATAAATAAGAAATTTCTGATGAGAGGTTTCAG 2100  
 QY 2101 TTTTTPAATCAACAGTCTTAAAGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2160

Db 2101 TTTTTPAATCAACAGTCTTAAAGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2160  
 QY 2161 TAAATACGTAAACCTATAAGAAATTAAGTTTATTAATTAAGGCAATTTATGCTGTGATAAT 2220  
 Db 2161 TAAATACGTAAACCTATAAGAAATTAAGTTTATTAATTAAGGCAATTTATGCTGTGATAAT 2220  
 QY 2221 TCTTACGGGAGAAAGAGATTTGATTTGGAAGAGCTTTGGGAAGAAAGTGTCTGCTGAAAT 2280  
 Db 2221 TCTTACGGGAGAAAGAGATTTGATTTGGAAGAGCTTTGGGAAGAAAGTGTCTGCTGAAAT 2280  
 QY 2281 TTCCAGAAATTTAAATTTGATTTGCTTACATAAACTTTTTCGACTTCAGAAAAAATAAAAA 2340  
 Db 2281 TTCCAGAAATTTAAATTTGATTTGCTTACATAAACTTTTTCGACTTCAGAAAAAATAAAAA 2340  
 QY 2341 AACAAAAAATAAAC 2354  
 Db 2341 AACAAAAAATAAAC 2354  
 RESULT 4  
 US-10-801-671-1  
 ; Sequence 1, Application US/10801671  
 ; Publication No. US20040152123A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: YAN, Chunhua et al.  
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
 ; FILE OF INVENTION: THEREOF  
 ; FILE REFERENCE: CL000758DIV-III  
 ; CURRENT APPLICATION NUMBER: US/10/801,671  
 ; PRIORITY FILING DATE: 2004-03-17  
 ; PRIOR APPLICATION NUMBER: 60/227,470  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: 09/810,671  
 ; PRIOR FILING DATE: 2001-03-19  
 ; NUMBER OF SEQ. ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2354  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-801-671-1  
 Query Match 100.0%; Score 2354; DB 18; Length 2354;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCCAGCTGGGGTTACTTTTAAAAAACAATGCTCCATGTCATCCCTCTTGAAGCTTCGCACT 60  
 Db 1 GCCAGCTGGGGTTACTTTTAAAAAACAATGCTCCATGTCATCCCTCTTGAAGCTTCGCACT 60  
 QY 61 CTGTTGAAGAGGACACTCATCCAGTCAATTTATTTAGAGCAAGGTCCTTTGAATGAGCGAG 120  
 Db 61 CTGTTGAAGAGGACACTCATCCAGTCAATTTATTTAGAGCAAGGTCCTTTGAATGAGCGAG 120  
 QY 121 ATTATCGGGAACGGAGATACGTTTGAACGAATACAGGAATGACTACTGTGAAGGATATGTC 180  
 Db 121 ATTATCGGGAACGGAGATACGTTTGAACGAATACAGGAATGACTACTGTGAAGGATATGTC 180  
 QY 181 CTAGACATTTACAGAGACATTTGAAAGCGGGTATCGAATCCACTGTCAGTAAATCTTCAG 240  
 Db 181 CTAGACATTTACAGAGACATTTGAAAGCGGGTATCGAATCCACTGTCAGTAAATCTTCAG 240  
 QY 241 TCCGACAGGAGAGAGCAGTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCAATCAGT 300  
 Db 241 TCCGACAGGAGAGAGCAGTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCAATCAGT 300  
 QY 301 CACGTTTGAAGAGCCACCGAGGAAAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTC 360  
 Db 301 CACGTTTGAAGAGCCACCGAGGAAAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTC 360  
 QY 361 ACCTGATCTGTCAAAGTGGAGACGTTTCTTAAGAGCAAGATATGAATCGTGACACTTTGG 420





APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 26852  
; LENGTH: 4035  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB4654-056-G4\_FLI  
US-10-425-114-26852

|                       |      |  |               |            |                      |
|-----------------------|------|--|---------------|------------|----------------------|
| Query Match           |      | 81.1%  | Score 1908.4; | DB 17;     | Length 4035;         |
| Best Local Similarity |      | 99.9%  | Pred. No. 0;  |            |                      |
| Matches 1909;         |      | Conservative   | 0;            | Mismatches | 1; Indels 0; Gaps 0; |
| Qy                    | 398  | ATATGAATCGTGGACACTTTGGGTGAAGAGCCCTTTGGCAAAAGTGTGAGAGTGCATTTGA  | 457           |            |                      |
| Db                    | 2126 | AGATGAATCGTGGACACTTTGGGTGAAGAGCCCTTTGGCAAAAGTGTGAGAGTGCATTTGA  | 2185          |            |                      |
| Qy                    | 458  | TCATGGCATGATGGCATGCAATGTAGCAGTGAATAATCGTAAAAATGATGAGCCGTTACCG  | 517           |            |                      |
| Db                    | 2186 | TCATGGCATGATGGCATGCAATGTAGCAGTGAATAATCGTAAAAATGATGAGCCGTTACCG  | 2245          |            |                      |
| Qy                    | 518  | TGAAGCAGCTCGTTCAGAAATCCCAAGTATTAGACACCTTAATAGTACTGATCCCAATAG   | 577           |            |                      |
| Db                    | 2246 | TGAAGCAGCTCGTTCAGAAATCCCAAGTATTAGACACCTTAATAGTACTGATCCCAATAG   | 2305          |            |                      |
| Qy                    | 578  | TGCTCTCCGATGTGTCAGATGCTAGAAATGGTTTGATCATCATGCTCATGTTGTGTTATGT  | 637           |            |                      |
| Db                    | 2306 | TGCTCTCCGATGTGTCAGATGCTAGAAATGGTTTGATCATCATGCTCATGTTGTGTTATGT  | 2365          |            |                      |
| Qy                    | 638  | GTTTGAACACTCTGGGACTTAGTACTTACGATTTCAATTAAGAAACAGCTTTCTGCCATT   | 697           |            |                      |
| Db                    | 2366 | GTTTGAACACTCTGGGACTTAGTACTTACGATTTCAATTAAGAAACAGCTTTCTGCCATT   | 2425          |            |                      |
| Qy                    | 698  | TCAATATGACACATCAGCAGATGGCGTATCAGATCTGCCAGTCAATTAATTTTACATA     | 757           |            |                      |
| Db                    | 2426 | TCAATATGACACATCAGCAGATGGCGTATCAGATCTGCCAGTCAATTAATTTTACATA     | 2485          |            |                      |
| Qy                    | 758  | TCATAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGTGTAAGTCTGA       | 817           |            |                      |
| Db                    | 2486 | TCATAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGTGTAAGTCTGA       | 2545          |            |                      |
| Qy                    | 818  | CTATGTAGTCAAAATATAATTTCTAAATGAACGCTGATGAACGCACACTGAAAAACACAGA  | 877           |            |                      |
| Db                    | 2546 | CTATGTAGTCAAAATATAATTTCTAAATGAACGCTGATGAACGCACACTGAAAAACACAGA  | 2605          |            |                      |
| Qy                    | 878  | TATCAAAAGTTGTGACTTTGGAAGTGCAACGTATGATGATGAACATCAGTACTTTGGT     | 937           |            |                      |
| Db                    | 2606 | TATCAAAAGTTGTGACTTTGGAAGTGCAACGTATGATGATGAACATCAGTACTTTGGT     | 2665          |            |                      |
| Qy                    | 938  | GTCTACCCGGCACTACAGAGTCTCCGAGTCAATTTGGCTTTAGTGTGCTCAGCCCTTG     | 997           |            |                      |
| Db                    | 2666 | GTCTACCCGGCACTACAGAGTCTCCGAGTCAATTTGGCTTTAGTGTGCTCAGCCCTTG     | 2725          |            |                      |
| Qy                    | 998  | TGATGTTTGGAGCATPAGGTTGCAATTTTATTGAATATTACCTTGGTTTTCACAGCTTTTCA | 1057          |            |                      |
| Db                    | 2726 | TGATGTTTGGAGCATPAGGTTGCAATTTTATTGAATATTACCTTGGTTTTCACAGCTTTTCA | 2785          |            |                      |
| Qy                    | 1058 | GACTCATGATAGTAAGACACCTGGCAATGATGGAACGAATATTAGACCCATACCA        | 1117          |            |                      |
| Db                    | 2786 | GACTCATGATAGTAAGACACCTGGCAATGATGGAACGAATATTAGACCCATACCA        | 2845          |            |                      |
| Qy                    | 1118 | ACACATGATTACAGAAAACAGAAAACGCAAGTATTTTCCACATACAGTAGATTGGGA      | 1177          |            |                      |
| Db                    | 2846 | ACACATGATTACAGAAAACAGAAAACGCAAGTATTTTCCACATACAGTAGATTGGGA      | 2905          |            |                      |
| Qy                    | 1178 | TGAACACAGTTCTGCTGCTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAAGGAATTTAT | 1237          |            |                      |

|    |      |  |      |  |  |
|----|------|--|------|--|--|
| Db | 2906 | TGAACACAGATTCTGCTGCTAGATATGTTAGGAGAGCGTGCMAACCGTTGAAGGAATTTAT            | 2965 |  |  |
| Qy | 1238 | GCTTTGTGTCATGATGAAGAACATGAGAACTGTTTGACCTGTTGCGAAGAAATGTTAGAATA           | 1297 |  |  |
| Db | 2966 | GCTTTGTGTCATGATGAAGAACATGAGAACTGTTTGACCTGTTGCGAAGAAATGTTAGAATA           | 3025 |  |  |
| Qy | 1298 | TGATCCAACTCAAGAAATTTACCTTTGGATGAAGCAATTCAGCATCTTTCTTTGACTTATTT           | 1357 |  |  |
| Db | 3026 | TGATCCAACTCAAGAAATTTACCTTTGGATGAAGCAATTCAGCATCTTTCTTTGACTTATTT           | 3085 |  |  |
| Qy | 1358 | AAAAAGAAATGAAATCGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGATTTACT               | 1417 |  |  |
| Db | 3086 | AAAAAGAAATGAAATCGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGATTTACT               | 3145 |  |  |
| Qy | 1418 | TAAGACTGTGTGCTCAACTCAATCTAATATTTTGTAAACATTAATAATTTTGTGA                  | 1477 |  |  |
| Db | 3146 | TAAGACTGTGTGCTCAACTCAATCTAATATTTTGTAAACATTAATAATTTTGTGA                  | 3205 |  |  |
| Qy | 1478 | CAGTTAAGTGAATATTTGTTATGTTTGTATCAATAGCATTAATTAACCTTGTAGCAAGT              | 1537 |  |  |
| Db | 3206 | CAGTTAAGTGAATATTTGTTATGTTTGTATCAATAGCATTAATTAACCTTGTAGCAAGT              | 3265 |  |  |
| Qy | 1538 | ATGCTCTTGTGATGCAATTAAGAAATTTAAATTTAAATTTTCTTTTGAATTTACCATTT              | 1597 |  |  |
| Db | 3266 | ATGCTCTTGTGATGCAATTAAGAAATTTAAATTTAAATTTTCTTTTGAATTTACCATTT              | 3325 |  |  |
| Qy | 1598 | TTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATAATTTAAATTTTCTTTTGAATTTACCATTT       | 1657 |  |  |
| Db | 3326 | TTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATAATTTAAATTTTCTTTTGAATTTACCATTT       | 3385 |  |  |
| Qy | 1658 | TACATGAGGTCACCTCTGAAAGTGAATTTTGTGTCAGTGAATAATTTAAATTTTCTTTTGAATTTACCATTT | 1717 |  |  |
| Db | 3386 | TACATGAGGTCACCTCTGAAAGTGAATTTTGTGTCAGTGAATAATTTAAATTTTCTTTTGAATTTACCATTT | 3445 |  |  |
| Qy | 1718 | TATTTCTTAAAGGAATATTTCTTTATATCTTCAATTTTGAATTTTAACTTTTAAAGTTTT             | 1777 |  |  |
| Db | 3446 | TATTTCTTAAAGGAATATTTCTTTATATCTTCAATTTTGAATTTTAACTTTTAAAGTTTT             | 3505 |  |  |
| Qy | 1778 | CTTCTGTAATTTGTTGAACGGGTGATTTATTTAACTCTAGATAAGCAGGTACTAGAAACC             | 1837 |  |  |
| Db | 3506 | CTTCTGTAATTTGTTGAACGGGTGATTTATTTAACTCTAGATAAGCAGGTACTAGAAACC             | 3565 |  |  |
| Qy | 1838 | AAAACTCAGAAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGTGTGTTCTTTT               | 1897 |  |  |
| Db | 3566 | AAAACTCAGAAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGTGTGTTCTTTT               | 3625 |  |  |
| Qy | 1898 | CATTGGGTGATGTCAGGGTGATAACAGCAATTCATGAAAGGATGCAAGTTTGTCCATT               | 1957 |  |  |
| Db | 3626 | CATTGGGTGATGTCAGGGTGATAACAGCAATTCATGAAAGGATGCAAGTTTGTCCATT               | 3685 |  |  |
| Qy | 1958 | GTGACAGTTTGTGTTTAAATAAAACACATACACACTTTATTTTAAAGATTTAAATCTA               | 2017 |  |  |
| Db | 3686 | GTGACAGTTTGTGTTTAAATAAAACACATACACACTTTATTTTAAAGATTTAAATCTA               | 3745 |  |  |
| Qy | 2018 | AAGTCAGTTTGGAAAATGGACATTTCCAAAGTATGTTTGGTGAAGTCAAGATATAAATA              | 2077 |  |  |
| Db | 3746 | AAGTCAGTTTGGAAAATGGACATTTCCAAAGTATGTTTGGTGAAGTCAAGATATAAATA              | 3805 |  |  |
| Qy | 2078 | GAAATTTCTGATGAGAGTTTCAAGTTTAAATCCCAAGTCTTACAGGCTTTAAACATTCGC             | 2137 |  |  |
| Db | 3806 | GAAATTTCTGATGAGAGTTTCAAGTTTAAATCCCAAGTCTTACAGGCTTTAAACATTCGC             | 3865 |  |  |
| Qy | 2138 | CAGCATCTGTTTATCAATGACATAAATACCTTAAACCTTATTAAGAAATTAAGTTTATTAAT           | 2197 |  |  |
| Db | 3866 | CAGCATCTGTTTATCAATGACATAAATACCTTAAACCTTATTAAGAAATTAAGTTTATTAAT           | 3925 |  |  |
| Qy | 2198 | AGGCAATTTATGCTGTGATTAATTTCTTACGGGAGAAAGAGATTTGATTTGAAAGCAGTT             | 2257 |  |  |
| Db | 3926 | AGGCAATTTATGCTGTGATTAATTTCTTACGGGAGAAAGAGATTTGATTTGAAAGCAGTT             | 3985 |  |  |
| Qy | 2258 | TGGGAAGAAAGTCTGCTGAAATTTTCCAGAAATTTAAATTTGATTTGGTTACAT                   | 2307 |  |  |
| Db | 3986 | TGGGAAGAAAGTCTGCTGAAATTTTCCAGAAATTTAAATTTGATTTGGTTACAT                   | 4035 |  |  |

RESULT 6

US-10-267-502-135  
 ; Sequence 135, Application US/10267502  
 ; Publication No. US20040071700A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kim, Jaeseob  
 ; TITLE OF INVENTION: Obesity Linked Genes  
 ; FILE REFERENCE: LSD-07416  
 ; CURRENT APPLICATION NUMBER: US/10/267,502  
 ; CURRENT FILING DATE: 2003-01-27  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 135  
 ; LENGTH: 1446  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-267-502-135

Query Match 54.6%; Score 1286; DB 17; Length 1446;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-255;  
 Matches 1286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |   |     |
|----|-----|---|-----|
| Qy | 85  | GTCTATTATTAGAACGAGGTCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG     | 144 |
| Db | 161 | GTCTATTATTAGAACGAGGTCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG     | 220 |
| Qy | 145 | ACGAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATATACAGACATG          | 204 |
| Db | 221 | ACGAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATATATACAGACATG        | 280 |
| Qy | 205 | AAAGCGGTATCGAATCCACATCGAGTAATCTTCAAGTCCGAGGAGGAGGAGTCTTA        | 264 |
| Db | 281 | AAAGCGGTATCGAATCCACATCGAGTAATCTTCAAGTCCGAGGAGGAGGAGTCTTA        | 340 |
| Qy | 265 | AAAGGAGCGCAATAGACATCTGTTCAAGTCAATCATGTCAGTTCGAAAGAGCCACCGAAGGA  | 324 |
| Db | 341 | AAAGGAGCGCAATAGACATCTGTTCAAGTCAATCATGTCAGTTCGAAAGAGCCACCGAAGGA  | 400 |
| Qy | 325 | AAAGTCCAGAGATATAGAGATGATGAGGAGGTCACCTGATCTGTCAAGTGGAGAGC        | 384 |
| Db | 401 | AAAGTCCAGAGATATAGAGATGATGAGGAGGTCACCTGATCTGTCAAGTGGAGAGC        | 460 |
| Qy | 385 | TTCTAAGAGCAAGATATGAATCTGGACACTTTGGGTGAAGGAGCCTTTGGCAAGTTG       | 444 |
| Db | 461 | TTCTAAGAGCAAGATATGAATCTGGACACTTTGGGTGAAGGAGCCTTTGGCAAGTTG       | 520 |
| Qy | 445 | TAGAGTGCAATGATCATGGCATGGATGCGCATGTCATGATGAGTGAAGTAAATCGTAAAGATG | 504 |
| Db | 521 | TAGAGTGCAATGATCATGGCATGGATGCGCATGTCATGATGAGTGAAGTAAATCGTAAAGATG | 580 |
| Qy | 505 | TAGCCGTTACCGTGAAGCAGCTGTTCCAGAAATCCAAATTTAGAGCAGCTTAAATAGTA     | 564 |
| Db | 581 | TAGCCGTTACCGTGAAGCAGCTGTTCCAGAAATCCAAATTTAGAGCAGCTTAAATAGTA     | 640 |
| Qy | 565 | CTGATCCCAATGATGTCCTCCGATGTGTCAGATGTCAGATGTCATCATCATGTC          | 624 |
| Db | 641 | CTGATCCCAATGATGTCCTCCGATGTGTCAGATGTCAGATGTCATCATCATGTC          | 700 |
| Qy | 625 | ATGTTTGTATGTTGTTGAATCTAGGACCTAGTACTTACGATTTTCAATTAAGAAAAACA     | 684 |
| Db | 701 | ATGTTTGTATGTTGTTGAATCTAGGACCTAGTACTTACGATTTTCAATTAAGAAAAACA     | 760 |
| Qy | 685 | GCCTTTCGCAATTTCAATTTGACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA      | 744 |
| Db | 761 | GCCTTTCGCAATTTCAATTTGACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA      | 820 |
| Qy | 745 | TAAATTTTACATCAATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTTGT       | 804 |
| Db | 821 | TAAATTTTACATCAATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTTGT       | 880 |

|    |      |   |      |
|----|------|---|------|
| Qy | 805  | TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCAAAATGAAACGTCATGAACGCACAC   | 864  |
| Db | 881  | TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCAAAATGAAACGTCATGAACGCACAC   | 940  |
| Qy | 865  | TGAAAAACACAGATATCAAAAGTTGTTGACCTTTGGAAGTGAACGTCATGATGAATCAATC   | 924  |
| Db | 941  | TGAAAAACACAGATATCAAAAGTTGTTGACCTTTGGAAGTGAACGTCATGATGAATCAATC   | 1000 |
| Qy | 925  | ACAGTACTTTGGTGTCTTACCCGGCCTACAGAGTCCCGAGGTCATTTGGCTTTAGGTT      | 984  |
| Db | 1001 | ACAGTACTTTGGTGTCTTACCCGGCCTACAGAGTCCCGAGGTCATTTGGCTTTAGGTT      | 1060 |
| Qy | 985  | GGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGTCATTTCTTATGAATATATACCTTGGT  | 1044 |
| Db | 1061 | GGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGTCATTTCTTATGAATATATACCTTGGT  | 1120 |
| Qy | 1045 | TCACAGTCTTTCAGACTCATGATAGTAAAGAGACCTGGCAATGATGAACGAATATAG       | 1104 |
| Db | 1121 | TCACAGTCTTTCAGACTCATGATAGTAAAGAGACCTGGCAATGATGAACGAATATAG       | 1180 |
| Qy | 1105 | GACCATACACACACATGATTTTCAGAAACAGAAACGCAAGTATTTTCACCATACAC        | 1164 |
| Db | 1181 | GACCATACACACACATGATTTTCAGAAACAGAAACGCAAGTATTTTCACCATACAC        | 1240 |
| Qy | 1165 | AGCTAGATTTGGGATGAACACAGTTCTCTGCTGTAGATATGTTAGGAGACGCTGCAAAACCGT | 1224 |
| Db | 1241 | AGCTAGATTTGGGATGAACACAGTTCTCTGCTGTAGATATGTTAGGAGACGCTGCAAAACCGT | 1300 |
| Qy | 1225 | TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGACCTGGTTGCGAA    | 1284 |
| Db | 1301 | TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGACCTGGTTGCGAA    | 1360 |
| Qy | 1285 | GAATGTTAGAAATGATGTCACAACTCAAGAAATTTACCTTGGATGAAGCAATTCAGCATCTT  | 1344 |
| Db | 1361 | GAATGTTAGAAATGATGTCACAACTCAAGAAATTTACCTTGGATGAAGCAATTCAGCATCTT  | 1420 |
| Qy | 1345 | TCCTTGACTTTATTAAGAAAGAAATGA                                     | 1370 |
| Db | 1421 | TCCTTGACTTTATTAAGAAAGAAATGA                                     | 1446 |

RESULT 7

US-10-641-643-699  
 ; Sequence 699, Application US/10641643  
 ; Publication No. US20040077003A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cocks, Benjamin G.  
 ; Jeffrey J. Seilhamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL  
 ; GENE EXPRESSION  
 ; NUMBER OF SEQUENCES: 1508  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/641,643  
 ; FILING DATE: 14-Aug-2003  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: <Unknown>  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Zeller, Karen J.





```
Db 779 ATGTTTGTATTGTTGAGCTGCTGGGACCTTAGTACCTATGATTTTATTAAAGAAATA 838
Qy 685 GCTTTCTGCCATTTCAAATTTGACCAATCAGGACATGCGGTATCAGATCTGCCAGTCAA 744
Db 839 GTTTTCTGCCATTTCAAATTTGATACATCAGGCAATGGCTTATCAGATCTGCCAGTCTA 898
Qy 745 TAAATTTTACATCATATAATAATTAACCCATACAGATCTGAGGCTGAAATATATTTTGT 804
Db 899 TAAATTTTACATCATATAATAATTAACACACAGGACCTTAAACCTGAAATATATTTTAT 958
Qy 805 TTGTGAAGCTGACATGATGAGTCAATATAATTTCTAAATGAAACGTCATGAACGACAC 864
Db 959 TTGTGAAGCTGACATGATGAGTCAATATAATTTCTAAATGAAACGATGAGCGCAT 1018
Qy 865 TGAATAACACAGATATCAAAGTTGTTGACTTTTGGAAAGTGCACGATGATGATGAACATC 924
Db 1019 TGAATAACACAGATATCAAAGTTGTTGATTTTGGAAAGTGCACATATGACGACGACATC 1078
Qy 925 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 984
Db 1079 ATAGTACTTTGGTGTCTACCAAGGCACTACAGGCTCCAGAGGTCAATTTTGGCTCTAGGTT 1138
Qy 985 GGTCTCAGCCTTTGATGTTTGGAGCATAGTTGCAATTTCTTAAATATTAACCTTTGGTT 1044
Db 1139 GGTCTCAGCCTTTGATGTTTGGAGCATAGGCTGCAATTTCTTAAATATTAACCTTTGGTT 1198
Qy 1045 TCACAGTCTTTGAGACTCATGATAGTAAAGAGCCTCGCAATGATGAAACGAAATATTAG 1104
Db 1199 TCACAGTCTTTGAGACTCATGATAGTAAAGAGCCTCGCAATGATGAAAGCGGATCTTAG 1258
Qy 1105 GACCCATACCAACACATGATTTCAAGAAACGCAAGTATTTTCAACCAATACC 1164
Db 1259 GACCCATACCAACATATGATTTCAAGAAACGCAAGTATTTTCAACCAATACC 1318
Qy 1165 AGCTAGATTGGATGAAACACAGTTCTGCTGTGTAGATATGTTAGGAGAGCGCTGCAACCGT 1224
Db 1319 AGCTAGATTGGAGCAGCATAGTTTCAAGTGGAGATATGTTAGGAGAGCGCTGCAACCGT 1378
Qy 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTCACTGGTTGCA 1284
Db 1379 TAAAGGAATTTATGCTTTGTCATGACGAAGAGCATGAGAGCTGTTGACCTGGTTGCA 1438
Qy 1285 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCGAGCATCCTT 1344
Db 1439 GAATGTTGAGTATGACCCAGGAGAGATCACCTTGGATGAAGCAATTCGAGCATCCTT 1498
Qy 1345 TCTTTGACTTATTAAGAAAGAAATGAAATGG 1375
Db 1499 TCTTTGACTTATTAAGAAAGAAATGAGTGGG 1529
```

## RESULT 10

US-10-267-502-138

; Sequence:138, Application US/10267502

; Publication No. US20040071700A1

; GENERAL INFORMATION:

; APPLICANT: Kim, Jaeseob

; APPLICANT: Galant, Ron

; TITLE OF INVENTION: Obesity Linked Genes

; FILE REFERENCE: LSD-07416

; CURRENT APPLICATION NUMBER: US/10/267,502

; CURRENT FILING DATE: 2003-01-27

; NUMBER OF SEQ ID NOS: 439

; SOFTWARE: Patencin version 3.2

; SEQ ID NO 138

; LENGTH: 1446

; TYPE: DNA

; ORGANISM: Mus musculus

US-10-267-502-138

Query Match

Best Local Similarity 46.6%; Score 1097.2; DB 17; Length 1446;

90.8%; Pred. No. 1.8e-216;

|         | Matches 1168;   | Conservative 0; | Mismatches 118; | Indels 0; | Gaps 0; |
|---------|---|-----------------|-----------------|-----------|---------|
| Qy 85   | GTCAATTTTAAAGCAAGTCTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG       | 144             |                 |           |         |
| Db 161  | GTCAATTTTAAAGCAAGTCTCTTGAATGACGAGATTTATCGGACCGGAGATACATTTG      | 220             |                 |           |         |
| Qy 145  | ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTATCACAGAGCATTTG    | 204             |                 |           |         |
| Db 221  | ATGAATACAGGAATGACTACTGTGAAGATATGTTCTCAAGACATTTACCATAGAGACGTTG   | 280             |                 |           |         |
| Qy 205  | AAAGCGGTATTCGAATCCACTGCGAGTAAATCTTCAGTCCGACGAGGAGAGAGCAGTCTTA   | 264             |                 |           |         |
| Db 281  | AAAGCACTTACCGGATCCATTTGCAATTAATCTCAGTCAGGAGGAGAGAGCAGCCTTA      | 340             |                 |           |         |
| Qy 265  | AAAGGAAGCGCAATAGACATCTGTTCAAGTCACTAGTCACTGTTGGAAGGACCCAGGA      | 324             |                 |           |         |
| Db 341  | AGAAAGCGTAAATAGACCTGTCAGTCACTAGTCACTGTCGAAGGACCCAGGAAGGA        | 400             |                 |           |         |
| Qy 325  | AAAGTCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGCG       | 384             |                 |           |         |
| Db 401  | AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGCG      | 460             |                 |           |         |
| Qy 385  | TTCTAAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCTTTGCAAGTTG    | 444             |                 |           |         |
| Db 461  | TTCTAAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCTTTGCAAGTTG    | 520             |                 |           |         |
| Qy 445  | TAGAGTGCATTTGATGATGCGCATGATGCGCATGATGATGAGTGAAGGAGGAGGAGG       | 504             |                 |           |         |
| Db 521  | TAGAGTGCATTTGATGATGCGCATGATGCGCATGATGATGAGTGAAGGAGGAGGAGG       | 580             |                 |           |         |
| Qy 505  | TAGGCGGTATTCGTTGAAGCAGTCTGTTCAAGAAATCCAAAGTATTTAGAGCAGTCTTA     | 564             |                 |           |         |
| Db 581  | TAGGACGTTTACCGGAGGAGCAGTCTGTTCTGAAATCCAAAGTATTTAGAGCAGTCTTA     | 640             |                 |           |         |
| Qy 565  | CTGATCCCAATAGTGTCTTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG       | 624             |                 |           |         |
| Db 641  | CTGACCCCAACAGTGTCTTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG       | 700             |                 |           |         |
| Qy 625  | ATGTTTGTATTTGTTGAACTACTGCGGACCTTAGTACTTACGATTTTCAATTAAGAAACA    | 684             |                 |           |         |
| Db 701  | ATGTTTGTATTTGTTGAACTACTGCGGACCTTAGTACTTACGATTTTCAATTAAGAAACA    | 760             |                 |           |         |
| Qy 685  | GCTTTCTGCCATTTCAAATTTGACCAATGACCAATCAGGAGATGCGGTATCAGATCTGCC    | 744             |                 |           |         |
| Db 761  | GTTTCTGCCATTTCAAATTTGATCAGATCAGGCAATGCGCTTATCAGATCTGCCAGTCTA    | 820             |                 |           |         |
| Qy 745  | TAAATTTTACATCATATAATAATTAACCCATACAGATCTGAGGCTGAAATATTTTGT       | 804             |                 |           |         |
| Db 821  | TAAATTTTACATCATATAATAATTAACCCATACAGATCTGAGGCTGAAATATTTTGT       | 880             |                 |           |         |
| Qy 805  | TTGTGAAGCTGACATGATGATCAATATAATTTCTAAATGAAACGTCATGATGAACGACAC    | 864             |                 |           |         |
| Db 881  | TTGTGAAGCTGACATGATGATCAATATAATTTCTAAATGAAACGTCATGATGAACGACAC    | 940             |                 |           |         |
| Qy 865  | TGAATAACACAGATATCAAAGTTGTTGACTTTTGGAAAGTGCACGATGATGATGAACATC    | 924             |                 |           |         |
| Db 941  | TGAATAACACAGATATCAAAGTTGTTGACTTTTGGAAAGTGCACATATGACGAGCAATC     | 1000            |                 |           |         |
| Qy 925  | ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT    | 984             |                 |           |         |
| Db 1001 | ATAGTACTTTGGTGTCTACCAAGGCACTACAGGCTTCCAGAGGTCATTTTGGCTCTAGGTT   | 1060            |                 |           |         |
| Qy 985  | GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATTTCTTAAATATTAACCTTTGGTT    | 1044            |                 |           |         |
| Db 1061 | GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATTTCTTAAATATTAACCTTTGGTT    | 1120            |                 |           |         |
| Qy 1045 | TCACAGTCTTTGAGACTCATGATAGTAAAGAGCAGCCTGCGCAATGATGAAACGAAATATTAG | 1104            |                 |           |         |
| Db 1121 | TCACAGTCTTTGAGACTCATGATAGTAAAGAGCAGCCTGCGCAATGATGAAACGAAATATTAG | 1180            |                 |           |         |
| Qy 1105 | GACCCATACCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG    | 1164            |                 |           |         |
| Db 1181 | GACCCATACCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG    | 1240            |                 |           |         |

[illegible]

```

RESULT 11
US-09-810-671-3
; Sequence 3, Application US/09810671
; Publication No. US20020076783A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-3

```

|    |       |   |       |
|----|-------|---|-------|
| Qy | 1708  | GACTACTTTATATCTCTAAAGGAATATCTCTTTATATACCTTCAAAATTTAGAACTTAACTTT   | 1767  |
| Db | 18618 | GACTACTTTATATCTCTAAAGGAATATCTCTTTATATACCTTCAAAATTTAGAACTTAACTTT   | 18677 |
| Qy | 1768  | AAAAGTTTTTCTCTCTAAATGTGTGAACGGGTGATTTATTATTAACCTCTAGATAAGCAGGT    | 1827  |
| Db | 18678 | AAAAGTTTTTCTCTCTAAATGTGTGAACGGGTGATTTATTATTAACCTCTAGATAAGCAGGT    | 18737 |
| Qy | 1828  | ACTAGAAACCAAACCTCAGAAAAATGTTTACTGTTAGAATTCCTATTPAATAATTTTAAGTGTTG | 1887  |
| Db | 18738 | ACTAGAAACCAAACCTCAGAAAAATGTTTACTGTTAGAATTCCTATTPAATAATTTTAAGTGTTG | 18797 |
| Qy | 1888  | TATTCCTTTTTTCATTTGGTGATGTCAGGGTGATAACGACACATTCATGAAAGGCATGCAG     | 1947  |
| Db | 18798 | TATTCCTTTTTTCATTTGGTGATGTCAGGGTGATAACGACACATTCATGAAAGGCATGCAG     | 18857 |
| Qy | 1948  | TTTGTCCATTTGACACAGTTTGTTTTAATAAAACACATACACACTTTATTTAAGATTAAAA     | 2007  |
| Db | 18858 | TTTGTCCATTTGACACAGTTTGTTTTAATAAAACACATACACACTTTATTTAAGATTAAAA     | 18917 |
| Qy | 2008  | TCTAACTGGAAAGTCAGCTTGGAAATATGGACAATTTCCAAAGTATGTTTGGTGAGTCCACAGA  | 2067  |
| Db | 18918 | TCTAACTGGAAAGTCAGCTTGGAAATATGGACAATTTCCAAAGTATGTTTGGTGAGTCCACAGA  | 18977 |
| Qy | 2068  | TATAAAAATAGAAATTCCTCATGAGAGGTTTCAGTTTTTAATACCAAGTCCTTTAGGAGTCT    | 2127  |
| Db | 18978 | TATAAAAATAGAAATTCCTCATGAGAGGTTTCAGTTTTTAATACCAAGTCCITTAGAGTCT     | 19037 |
| Qy | 2128  | TAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAAGAAATTAAG    | 2187  |
| Db | 19038 | TAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAAGAAATTAAG    | 19097 |
| Qy | 2188  | TTTTATTAAATTAGCAATTTATGTCGTGTGATAATTTCTTACGGGAGAAAGGATTTGATTG     | 2247  |
| Db | 19098 | TTTTATTAAATTAGCAATTTATGTCGTGTGATAATTTCTTACGGGAGAAAGGATTTGATTG     | 19157 |
| Qy | 2248  | GAAGACAGTTTGGGAAGAAGTGCTGCTCAAAATTTCCAGAAATTTAAATGATTGGTTACAT     | 2307  |
| Db | 19158 | GAAGACAGTTTGGGAAGAAGTGCTGCTGAAATTTCCAGAAATTTAAATGATTGGTTACAT      | 19217 |
| Qy | 2308  | AAACTTTTTGACTTCAG 2324  |       |
| Db | 19218 | AAACTTTTTGACTTCAG 19234   |       |

```

RESULT 12
US-10-109-854-3
; Sequence 3, Application US/10109854
; Publication No. US20020119548A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: fastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-3

```

|                       |              |                     |               |               |
|-----------------------|--------------|---------------------|---------------|---------------|
| Query Match           | 46.6%;       | Score 1097;         | DB 13;        | Length 21234; |
| Best Local Similarity | 100.0%;      | Pred. No. 6.7e-216; |               |               |
| Matches 1097;         | Conservative | 0;                  | Mismatches 0; | Indels 0;     |
|                       |              |                     |               | Gaps 0;       |

```
QY 1228 AGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAAACTGTTGACCTGGTTCGAAGAA 1287
Db 18138 AGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAAACTGTTGACCTGGTTCGAAGAA 18197
QY 1288 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTCGAGCATCTTTCT 1347
Db 18198 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTCGAGCATCTTTCT 18257
QY 1348 TTGACTTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTTCTCTAGA 1407
Db 18258 TTGACTTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTTCTCTAGA 18317
QY 1408 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 1467
Db 18318 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 18377
QY 1468 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATATTAACCTTG 1527
Db 18378 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATATTAACCTTG 18437
QY 1528 TTAACCAAGTAGTGTCTTGAATAATGCATTTAGAAAAATTAATAATTTTCTTTTGA 1587
Db 18438 TTAACCAAGTAGTGTCTTGAATAATGCATTTAGAAAAATTAATAATTTTCTTTTGA 18497
QY 1588 ATTACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATATAATGTGATGTC 1647
Db 18498 ATTACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATATAATGTGATGTC 18557
QY 1648 TTGCCTTTTGTACATGGAGGTCACCTCTGAAGTGATTTTTTTTGTAGTAAAGGAATCTT 1707
Db 18558 TTGCCTTTTGTACATGGAGGTCACCTCTGAAGTGATTTTTTTTGTAGTAAAGGAATCTT 18617
QY 1708 GACTACTTTATATCTTAAAGGAATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT 1767
Db 18618 GACTACTTTATATCTTAAAGGAATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT 18677
QY 1768 AAAAGTTTTTCTCTGTAATTTGTGAACGGGTGATTTATTTAACTCTAGATAAGCAGGT 1827
Db 18678 AAAAGTTTTTCTCTGTAATTTGTGAACGGGTGATTTATTTAACTCTAGATAAGCAGGT 18737
QY 1828 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATCTTATTAATTTTAAAGTGTG 1887
Db 18738 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATCTTATTAATTTTAAAGTGTG 18797
QY 1888 TATCTCTTTTCAATTCGGGTGATGTCAGGGTGATTAACAGACATTCATGGAAGGCGATGAG 1947
Db 18798 TATCTCTTTTCAATTCGGGTGATGTCAGGGTGATTAACAGACATTCATGGAAGGCGATGAG 18857
QY 1948 TTTGTCCATTTGTGACAGTTTGTTTAATAAACCACATACACACTTTTATTAAGATTTAAAA 2007
Db 18858 TTTGTCCATTTGTGACAGTTTGTTTAATAAACCACATACACACTTTTATTAAGATTTAAAA 18917
QY 2008 TCTAACTGGAAGTCAAGCTTGGAAAAATGGACATTTCCAAAGTATGTTGGTGAGTCACAGA 2067
Db 18918 TCTAACTGGAAGTCAAGCTTGGAAAAATGGACATTTCCAAAGTATGTTGGTGAGTCACAGA 18977
QY 2068 TATAAATATAGAAATCTGTGATGAGAGGTTTCAGTTTATTAACCAAGTCCCTTAGAGTCT 2127
Db 18978 TATAAATATAGAAATCTGTGATGAGAGGTTTCAGTTTATTAACCAAGTCCCTTAGAGTCT 19037
QY 2128 TAAATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAAACCTTATAGAAATTAAG 2187
Db 19038 TAAATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAAACCTTATAGAAATTAAG 19097
QY 2188 TTTATTAATAGGCAATTTATGTCGTGATTAATCTTACGGGAGAAAGAGGATTTGATG 2247
Db 19098 TTTATTAATAGGCAATTTATGTCGTGATTAATCTTACGGGAGAAAGAGGATTTGATG 19157
QY 2248 GAAACGATTTGGGAGAAAGTGTGCTGAAATTTCCAGAAATTTAAATGATTTGGTTACAT 2307
Db 19158 GAAACGATTTGGGAGAAAGTGTGCTGAAATTTCCAGAAATTTAAATGATTTGGTTACAT 19217
```

```
QY 2308 AAACCTTTTGTGACTTCAG 2324
Db 19218 AAACCTTTTGTGACTTCAG 19234

RESULT 13
US-10-339-656-3
; Sequence 3, Application US/10339656
; Publication No. US20030134319A1
; GENERAL INFORMATION:
; APPLICANT: YAN Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3.
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-3

Query Match 46.6%; Score 1097; DB 15; Length 21234;
Best Local Similarity 100.0%; Pred. No. 6.7e-216;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1228 AGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAAACTGTTGACCTGGTTCGAAGAA 1287
Db 18138 AGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAAACTGTTGACCTGGTTCGAAGAA 18197
QY 1288 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTCGAGCATCTTTCT 1347
Db 18198 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTCGAGCATCTTTCT 18257
QY 1348 TTGACTTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTTCTCTAGA 1407
Db 18258 TTGACTTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTTCTCTAGA 18317
QY 1408 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 1467
Db 18318 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 18377
QY 1468 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATATTAACCTTG 1527
Db 18378 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATATTAACCTTG 18437
QY 1528 TTAACCAAGTAGTGTCTTGAATAATGCATTTAGAAAAATTAATAATTTTCTTTTGA 1587
Db 18438 TTAACCAAGTAGTGTCTTGAATAATGCATTTAGAAAAATTAATAATTTTCTTTTGA 18497
QY 1588 ATTACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATATAATGTGATGTC 1647
Db 18498 ATTACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATATAATGTGATGTC 18557
QY 1648 TTGCCTTTTGTACATGGAGGTCACCTCTGAAGTGATTTTTTTTGTAGTAAAGGAATCTT 1707
Db 18558 TTGCCTTTTGTACATGGAGGTCACCTCTGAAGTGATTTTTTTTGTAGTAAAGGAATCTT 18617
QY 1708 GACTACTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT 1767
Db 18618 GACTACTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT 18677
QY 1768 AAAAGTTTTTCTCTGTAATTTGTGAACGGGTGATTTATTTAACTCTAGATAAGCAGGT 1827
```







|                       |       |   |          |            |    |        |      |
|-----------------------|-------|---|----------|------------|----|--------|------|
| Query Match           | 41.5% | Score   | 977.4    | DB         | 17 | Length | 3040 |
| Best Local Similarity | 99.9% | Pred. No.   | 1.4e-191 |            |    |        |      |
| Matches               | 978   | Conservative  | 0        | Mismatches | 1  | Indels | 0    |
|                       |       |   |          |            |    |        |      |
| QY                    | 398   | ATATGAATCGTGGACACCTTTGGGTGAAGAGCGCTTTGGCAAAGTTGTAGAGTGCATTGA    | 457      |            |    |        |      |
| DB                    | 2062  | AGATGAATCGTGGACACCTTTGGGTGAAGAGCGCTTTGGCAAAGTTGTAGAGTGCATTGA    | 2121     |            |    |        |      |
| QY                    | 458   | TCATGGCATCGATGGCGATGATGTAGCAGTGAATAATCGTAAAAAATCTAGCGCGTTACCG   | 517      |            |    |        |      |
| DB                    | 2122  | TCATGGCATCGATGGCGATGATGTAGCAGTGAATAATCGTAAAAAATCTAGCGCGTTACCG   | 2181     |            |    |        |      |
| QY                    | 518   | TGAAGCAGCTCGTTTCAGAAAATCCAAAGTATTAGAGCACTTAAATAGTACTCATCCCAATAG | 577      |            |    |        |      |
| DB                    | 2182  | TGAAGCAGCTCGTTTCAGAAAATCCAAAGTATTAGAGCACTTAAATAGTACTCATCCCAATAG | 2241     |            |    |        |      |
| QY                    | 578   | TGTCTCCGATGTGTCAGATGCTAGAAATGGTTTGATCATCATGGTCATGTTTGTATGTT     | 637      |            |    |        |      |
| DB                    | 2242  | TGTCTCCGATGTGTCAGATGCTAGAAATGGTTTGATCATCATGGTCATGTTTGTATGTT     | 2301     |            |    |        |      |
| QY                    | 638   | GTTTGAACTACTGGGACTTGTACTTACGATTTTCATTAAGAAAAACAGCTTCTCGCCATT    | 697      |            |    |        |      |
| DB                    | 2302  | GTTTGAACTACTGGGACTTGTACTTACGATTTTCATTAAGAAAAACAGCTTCTCGCCATT    | 2361     |            |    |        |      |
| QY                    | 698   | TCAAATTTGACCATCAGGCAGATGGCGTATCAGATCTGCAGTCAATAAATTTTTTACA      | 757      |            |    |        |      |
| DB                    | 2362  | TCAAATTTGACCATCAGGCAGATGGCGTATCAGATCTGCAGTCAATAAATTTTTTACA      | 2421     |            |    |        |      |
| QY                    | 758   | TCATAATAAATTAAACCCATACAGATCTGAAAGCCCTGAAAAATATTTGTTGTTGAAGTCTGA | 817      |            |    |        |      |
| DB                    | 2422  | TCATAATAAATTAAACCCATACAGATCTGAAAGCCCTGAAAAATATTTGTTGTTGAAGTCTGA | 2481     |            |    |        |      |
| QY                    | 818   | CTATGTAGTCAAAATATAATTTCTAAAAATGAAACGATGAAACGACACTGAAAAACACAGA   | 877      |            |    |        |      |
| DB                    | 2482  | CTATGTAGTCAAAATATAATTTCTAAAAATGAAACGATGAAACGACACTGAAAAACACAGA   | 2541     |            |    |        |      |
| QY                    | 878   | TATCAAAGTTGTTGACTTTTGGAGTGCAAACGTATGATGATGAACATCACAGTACTTTGGT   | 937      |            |    |        |      |
| DB                    | 2542  | TATCAAAGTTGTTGACTTTTGGAGTGCAAACGTATGATGATGAACATCACAGTACTTTGGT   | 2601     |            |    |        |      |
| QY                    | 938   | GTCTACCCGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGTTGGTCTCAGCCCTTG      | 997      |            |    |        |      |
| DB                    | 2602  | GTCTACCCGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGTTGGTCTCAGCCCTTG      | 2661     |            |    |        |      |
| QY                    | 998   | TGATGTTTGGAGCATAGGTTGCATTTCTTATTGAATATATACCTTGGTTTCACAGTCTTTCA  | 1057     |            |    |        |      |
| DB                    | 2662  | TGATGTTTGGAGCATAGGTTGCATTTCTTATTGAATATATACCTTGGTTTCACAGTCTTTCA  | 2721     |            |    |        |      |
| QY                    | 1058  | GACTCATGATAGTAAAGAGCACCTGGCAATGATGGAAACGAATATTAGGAACCCATACACA   | 1117     |            |    |        |      |

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 23:50:16 ; Search time 166 Seconds  
(without alignments)  
1036.797 Million cell updates/sec

Title: US-10-801-671-2

Perfect score: 2410  
Sequence: 1 MCIPLEASHSVEDTHPSHY.....QRITLDEALQHPPDLLKKK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 2410   | 100.0       | 445    | 5     | Aae20170 Human kin |
| 2          | 2315.5 | 96.1        | 481    | 8     | ADI26143 Human pro |
| 3          | 2315.5 | 96.1        | 481    | 8     | ADO08026 Human pol |
| 4          | 2315.5 | 96.1        | 481    | 8     | ADP24394 PRO polyp |
| 5          | 2294.5 | 95.2        | 481    | 4     | Aab69705 Human cel |
| 6          | 2261.5 | 93.8        | 481    | 8     | ADI26141 Human pro |
| 7          | 2252.5 | 93.5        | 481    | 8     | ADO08029 Human pro |
| 8          | 2177.5 | 90.4        | 479    | 2     | AAW49914 Mouse CLK |
| 9          | 2127.5 | 88.3        | 499    | 4     | AAB65648 Novel pro |
| 10         | 2127.5 | 88.3        | 499    | 8     | ADI29255 Human MAR |
| 11         | 2117.5 | 87.9        | 451    | 2     | AAV27054 Human pro |
| 12         | 1893   | 78.5        | 484    | 8     | ADO08027 Human pol |
| 13         | 1887   | 78.3        | 484    | 5     | AAU75108 Serine an |
| 14         | 1887   | 78.3        | 484    | 8     | ADH58849 Human CDC |
| 15         | 1887   | 78.3        | 484    | 8     | ADR14001 Human NCF |
| 16         | 1887   | 78.3        | 484    | 8     | ADP24719 PRO polyp |
| 17         | 1823   | 75.6        | 483    | 8     | ADO08030 Mouse pol |
| 18         | 1808.5 | 75.0        | 482    | 2     | AAW49911 Mouse CLK |
| 19         | 1803   | 74.8        | 453    | 8     | ADI26145 Human pro |
| 20         | 1607   | 66.7        | 301    | 8     | ADO08031 Mouse pol |
| 21         | 1549   | 64.3        | 441    | 7     | ADG74657 Human kin |
| 22         | 1535   | 63.7        | 283    | 3     | AAB33778 Human sec |
| 23         | 1526   | 63.3        | 350    | 3     | AAB58831 Breast an |
| 24         | 1517   | 62.9        | 283    | 3     | AAB33777 Human sec |
| 25         | 1422.5 | 59.0        | 498    | 6     | Aae37967 Human kin |

## ALIGNMENTS

### RESULT 1

AAE20170

ID AAE20170 standard; protein; 445 AA.

XX AAE20170;

AC AAE20170;

XX 18-JUN-2002 (first entry)

XX Human kinase protein.

DE Human kinase protein.

XX Human; kinase protein; enzyme; cytosolic; osteopathic; gene expression;

KW colon-moderately differentiated adenocarcinoma; chromosome mapping; gene;

KW drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;

KW bone osteosarcoma.

XX Homo sapiens.

XX Key

FT Modified-site 10.13

FT Modified-site 24.31

FT Modified-site 25.28

FT Modified-site 29.36

FT Modified-site 55.61

FT Modified-site 69.71

FT Modified-site 72.74

FT Modified-site 73.76

FT Modified-site 76.78

FT Modified-site 94.96

FT Modified-site 97.100

FT Modified-site 102.105

FT Modified-site 128.131

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Binding-site 129.153

FT Modified-site 147..152 /note= "N-myristoylation site"  
FT Modified-site 176..179 /note= "Asn is N-glycosylated"  
FT Modified-site 209..212 /note= "Casein kinase II phosphorylation site"  
FT Active-site 246..258 /note= "Serine/threonine protein kinase active-site signature"  
FT Modified-site 247..250 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 277..279 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 292..295 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 303..305 /note= "Protein kinase C phosphorylation site"  
FT Region 324..344 /note= "Helix 1"  
FT Modified-site 368..370 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 425..427 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 429..432 /note= "Casein kinase II phosphorylation site"  
XX WO200216567-A2.  
XX 28-FEB-2002.  
XX 24-AUG-2001; 2001WO-US026389.  
XX 24-AUG-2000; 2000US-0227470P.  
XX 19-MAR-2001; 2001US-00810671.  
XX (APPL-) APPLERA CORP.  
XX Yan C, Ye J, Ketchum KA, Di Francesco V, Beasley EM;  
XX WPI; 2002-269354/31.  
XX N-PSDB; AAD32038.  
XX New human kinase proteins and nucleic acids, useful in drug screening  
XX assays, identifying modulators of kinase activity or treating disorders  
XX characterized by absence or unwanted expression of the protein.  
XX Claim 1; Fig 2; 8lpp; English.  
XX The invention relates to isolated human kinase proteins and nucleic  
XX acids. The nucleic acid and peptide sequences can be used as models for  
XX the development of human therapeutic targets, aid in the identification  
XX of therapeutic proteins and serve as targets for the development of human  
XX therapeutic agents that modulate kinase activity in cells and tissues  
XX that express the kinase. The nucleic acids are useful as probes or  
XX primers, in constructing recombinant vectors, for expressing antigenic  
XX portions of the proteins, chromosome mapping, drug screening, testing an  
XX individual for a genotype, and for gene therapy in patients containing  
XX cells that are aberrant in kinase gene expression. The proteins may be  
XX used in drug screening assays, in the identification of compounds that  
XX modulate, stimulate or inhibit kinase activity, in pharmacogenomic  
XX analysis, in treating disorders characterised by an absence or unwanted  
XX expression of the protein (bone osteosarcoma, or colon-moderately  
XX differentiated adenocarcinoma), and in generating antibodies specific for  
XX the peptides. Such antibodies can be used to detect the protein in situ,  
XX in vitro, or in cell lysate or supernatant, to isolate and purify the  
XX proteins from host cells, pharmacogenomic analysis, tissue typing, and in  
XX inhibiting protein function. The present sequence is human kinase  
XX protein. Human kinase protein gene is located on chromosome 5  
XX Sequence 445 AA;

Query Match 100.0%; Score 2410; DB 5; Length 445;  
Best Local Similarity 100.0%; Pred. No. 2.8e-221;

Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MCIPLEASHSVEDTHPSHYLEARSINERDYRRRVYVDEYRNDYCEGYVPRHYRDIESG 60  
DB 1 MCIPLEASHSVEDTHPSHYLEARSINERDYRRRVYVDEYRNDYCEGYVPRHYRDIESG 60  
QY 61 YRIHCSKSSVRSRRSSPKRKRNRHSCSHQSRKSRKSRSSIEDDEEGHLICQSGDVL 120  
DB 61 YRIHCSKSSVRSRRSSPKRKRNRHSCSHQSRKSRKSRSSIEDDEEGHLICQSGDVL 120  
QY 121 ARYEIVDTLGEAGFGKVVCEIDHGMDGMVAVKIVKNGRYREARSEIQVLEHLNSTDP 180  
DB 121 ARYEIVDTLGEAGFGKVVCEIDHGMDGMVAVKIVKNGRYREARSEIQVLEHLNSTDP 180  
QY 181 NSVFRVCQMLEWDFDHGHVCIVFELLGLSTYDFIKENSLPFOIDHIROMAYOICOSINF 240  
DB 181 NSVFRVCQMLEWDFDHGHVCIVFELLGLSTYDFIKENSLPFOIDHIROMAYOICOSINF 240  
QY 241 LHENKLTHTDLKPENILFVKSDYVVKYNSKMKRDETLKNTDIKVVDFSGSATYDDEHST 300  
DB 241 LHENKLTHTDLKPENILFVKSDYVVKYNSKMKRDETLKNTDIKVVDFSGSATYDDEHST 300  
QY 301 LVSTRHYRAPEVILALGWSQPCDVNSIGCTLIIEYILGFTVFQTHDSKEHLAMMERILGPI 360  
DB 301 LVSTRHYRAPEVILALGWSQPCDVNSIGCTLIIEYILGFTVFQTHDSKEHLAMMERILGPI 360  
QY 361 PQHMIQKTRKRYFHNQLDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRML 420  
DB 361 PQHMIQKTRKRYFHNQLDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRML 420  
QY 421 EYDPTQRITLDEALQHPFFDLKKK 445  
DB 421 EYDPTQRITLDEALQHPFFDLKKK 445  
RESULT 2  
ADI26143  
ID ADI26143 standard; protein; 481 AA.  
XX AC ADI26143;  
XX AC ADI26143;  
DT 22-APR-2004 (first entry)  
DE Human protein that promotes STAT6 activation #54.  
XX human; signal transducer and activator of transcription 6; STAT6;  
XX immunogen; STAT6 activation; allergy; inflammation; autoimmune disease;  
XX diabetes; hyperlipidaemia; infection; cancer; Th1 hyperactive disease;  
XX rheumatoid arthritis; osteoarthritis; systemic lupus erythematosus;  
XX sepsis; asthma; allergic rhinitis; ischaemic heart disease;  
XX subarachnoid haemorrhage; viral hepatitis; AIDS.  
XX Homo sapiens.  
XX WO2003104277-A2.  
XX 18-DEC-2003.  
XX 05-JUN-2003; 2003WO-JP007123.  
XX 05-JUN-2002; 2002JP-00164257.  
XX 06-JUN-2002; 2002US-0385912P.  
XX 26-DEC-2002; 2002JP-00377326.  
XX 27-DEC-2002; 2002US-0436467P.  
XX 15-MAY-2003; 2003JP-00137505.  
XX 16-MAY-2003; 2003US-0470836P.  
XX (ASAH ) ASAH KASEI KK.  
XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;  
XX WPI; 2004-122214/12.  
XX N-PSDB; ADI26142.

XX New signal transducer and activator of transcription 6 activation  
PT promoting purified protein, for diagnosing and treating disease  
PT associated with activation/inhibition of transcription factor e.g.  
PT diabetes and cancer.

PS Claim 1; SEQ ID NO 108; 1368pp; English.

XX The invention relates to a purified protein promoting signal transducer  
CC and activator of transcription 6 activation (STAT6). The protein is  
CC useful for the producing an antibody, which involves administering the  
CC protein or its epitope-bearing fragments to a non-human animal as an  
CC antigen. The nucleic acid is useful for diagnosing a disease or  
CC susceptibility to a disease related to expression or activity of the  
CC protein. A transformant expressing the protein is useful for screening  
CC compounds which inhibit or promote STAT6 activation. A transformant  
CC expressing the protein is useful for producing a pharmaceutical  
CC composition. Compositions, antibodies and antisense molecules are useful  
CC for the treating a disease associated with STAT6 activation such as  
CC allergic diseases, inflammation, autoimmune diseases, diabetes,  
CC hyperlipidaemia, infections disease and cancers. Compositions are useful  
CC for treating disease associated with STAT6 activation and/or prevention  
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid  
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,  
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,  
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6  
CC activity. The protein or nucleic acid is effectively useful for screening  
CC compounds for treating and preventing disease associated with excessive  
CC activation or inhibition of STAT6. The present sequence represents the  
CC amino acid sequence of a human protein which promotes STAT6 activation.

XX Sequence 481 AA;

Query Match 96.1%; Score 2315.5; DB 8; Length 481;

Best Local Similarity 96.2%; Pred. No. 3.5e-212; Indels 11; Gaps 2;  
Matches 432; Conservative 2; Mismatches 4;

Qy 8 SH-SVEEDTH-----PSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRD 56  
Db 33 SHSSTQENRHCKPHQPKESDCHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRD 92  
Qy 57 IESGVRHCKSSVRSRSPKRNHCHSCSHQSKSRHRRKRSIEDDEGHLCOSG 116  
Db 93 IESGVRHCKSSVRSRSPKRNHCHSCSHQSKSRHRRKRSIEDDEGHLCOSG 152  
Qy 117 DVLRYVEIVDTLGEAGFKVVECIDHGMGMHVAVKIVKNVGRYREARSEIQVLEHLN 176  
Db 153 DVLRYVEIVDTLGEAGFKVVECIDHGMGMHVAVKIVKNVGRYREARSEIQVLEHLN 212  
Qy 177 STDPSNVCQMLEWFDDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRQWAYQICQ 236  
Db 213 STDPSNVCQMLEWFDDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRQWAYQICQ 272  
Qy 237 SINFLHNKLTHTDLKPENILFKVSDYVKNVKNMKRDLTKNTDIKVPDFGSATYDE 296  
Db 273 SINFLHNKLTHTDLKPENILFKVSDYVKNVKNMKRDLTKNTDIKVPDFGSATYDE 332  
Qy 297 HHSTLVSTRHRAPEVILALGWSQCDVWSGTCILIEYLGFTVFTQTHDSKEHLAMMERI 356  
Db 333 HHSTLVSTRHRAPEVILALGWSQCDVWSGTCILIEYLGFTVFTQTHDSKEHLAMMERI 392  
Qy 357 LGPIPOHMIQTRKKYFHHNQLDWDEHSSAGRYVRRCKPKLKEPMLCHDEHEKLPDLV 416  
Db 393 LGPIPOHMIQTRKKYFHHNQLDWDEHSSAGRYVRRCKPKLKEPMLCHDEHEKLPDLV 452  
Qy 417 RRMLEYDPTQRTITLDEALQHPFFDLKKK 445  
Db 453 RRMLEYDPTQRTITLDEALQHPFFDLKKK 481

RESULT 3  
ADO08026  
ID ADO08026 standard; protein; 481 AA.

XX ADO08026;  
XX 01-JUL-2004 (first entry)  
XX Human polypeptide #39.  
XX Human; fat cell number; fat cell size; obesity; diabetes; anorectic;  
XX antidiabetic.  
XX Homo sapiens.  
XX US2004071700-A1.  
XX 15-APR-2004.  
XX 09-OCT-2002; 2002US-00267502.  
XX 09-OCT-2002; 2002US-00267502.  
XX (LIFE-) LIFE SCI DEV CORP.  
XX Kim J, Galant R;  
XX WPI; 2004-328526/30.  
XX N-ESDB; ADO07809.  
XX Identifying compounds that influence fat cell number or size for treating  
PT or preventing obesity or diabetes by exposing the cell to the agent and  
PT identifying fat cell number or size relative to cells not exposed to the  
PT agent.

PS Claim 14; SEQ ID NO 352; 275pp; English.

XX The invention relates to a method of identifying compounds that influence  
CC fat cell number or size comprising providing a cell that expresses a gene  
CC and an agent, exposing the cell to the agent and identifying fat cell  
CC number or size relative to cells not exposed to the agent. The method  
CC also comprises providing an expression vector and an agent, exposing the  
CC vector to the agent, detecting a change in expression of the gene  
CC relative to expression of the gene in an expression vector not exposed to  
CC the agent, treating a subject with the agent and identifying fat cell  
CC number or size in the subject. The agent comprises an antisense  
CC oligonucleotide. The subject comprises a mammal, preferably a human. The  
CC method also comprises providing a polypeptide and an agent, exposing the  
CC polypeptide to the agent, detecting binding of the agent to the  
CC polypeptide or a change in an activity of the polypeptide, treating a  
CC subject with the agent and identifying fat cell number or size in the  
CC subject. The agent comprises an antibody. A method of regulating fat cell  
CC number or size comprises providing a subject containing fat cells and an  
CC agent that changes the expression of a gene, and treating the subject  
CC with the agent under conditions so that fat cell size or number in the  
CC subject is altered. The method is useful for identifying compounds that  
CC influence fat cell number or size, for preparing a composition for  
CC treating or preventing obesity or diabetes. This sequence represents a  
CC human polypeptide used in the scope of the invention.

XX Sequence 481 AA;

Query Match 96.1%; Score 2315.5; DB 8; Length 481;

Best Local Similarity 96.2%; Pred. No. 3.5e-212;  
Matches 432; Conservative 2; Mismatches 4; Indels 11; Gaps 2;

Qy 8 SH-SVEEDTH-----PSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRD 56  
Db 33 SHSSTQENRHCKPHQPKESDCHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRD 92  
Qy 57 IESGVRHCKSSVRSRSPKRNHCHSCSHQSKSRHRRKRSIEDDEGHLCOSG 116  
Db 93 IESGVRHCKSSVRSRSPKRNHCHSCSHQSKSRHRRKRSIEDDEGHLCOSG 152  
Qy 117 DVLRYVEIVDTLGEAGFKVVECIDHGMGMHVAVKIVKNVGRYREARSEIQVLEHLN 176

Db 153 DVLARVEIVDTLGEAGFKVVEICDHGMGMHVAVKIVKNGRYREARSEIQVLEHLN 212  
 Qy 177 STDPSNFRVCVQMLEWDFHGHGVCIVFELLGLSYDFIKENSFLPFQIDHIRMAYQICQ 236  
 Db 213 STDPSNFRVCVQMLEWDFHGHGVCIVFELLGLSYDFIKENSFLPFQIDHIRMAYQICQ 272  
 Qy 237 SINFLHNKLTHDLPENILFVKSDDYVVKVNSKMKRDETLKNTDIKVVDFFGSATYDDE 296  
 Db 273 SINFLHNKLTHDLPENILFVKSDDYVVKVNSKMKRDETLKNTDIKVVDFFGSATYDDE 332  
 Qy 297 HSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYILGFTVFOHDSKEHLAMMERI 356  
 Db 333 HSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYILGFTVFOHDSKEHLAMMERI 392  
 Qy 357 LGPIPOHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLV 416  
 Db 393 LGPIPOHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLV 452  
 Qy 417 RRMLEYDPTORITLDEALQHPFFDLKKK 445  
 Db 453 RRMLEYDPTORITLDEALQHPFFDLKKK 481

RESULT 4  
 ADP24394  
 ID ADP24394 standard; protein; 481 AA.

AC ADP24394;  
 DT 18-NOV-2004 (first entry)  
 DE PRO polypeptide SEQ ID NO:1572.

KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;  
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;  
 KW antiaschmatic; hepatotropic; respiratory; gene therapy; immune system.  
 OS Unidentified.

PN WO2004041170-A2.  
 XX 21-MAY-2004.  
 XX 30-OCT-2003; 2003WO-US034312.  
 XX 01-NOV-2002; 2002US-0423394P.  
 XX (GETH ) GENENTECH INC.  
 XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WT;  
 PI Wu TD;  
 XX WPI; 2004-419628/39.  
 DR N-PSDB; ADP24393.

PT New PRO polypeptides and polynucleotides, useful for treating e.g.  
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
 PT renal disease, or demyelinating diseases of the central or peripheral  
 PT nervous system.

PS Claim 7; SEQ ID NO 1572; 2940pp; English.

XX The invention relates to a novel isolated nucleic acid and the PRO  
 CC polypeptide encoded by it. A protein of the invention has  
 CC antiinflammatory, antirheumatic, antipsoriatic, immunosuppressive,  
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
 CC antiaschmatic, hepatotropic, and respiratory activity. A polynucleotide  
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
 CC agonist, antagonist, or antibody that specifically binds to the  
 CC polypeptide is useful for treating an immune related disorder such as  
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic

CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
 CC disease, a demyelinating disease of the central or peripheral nervous  
 CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,  
 CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary  
 CC disease, infectious or autoimmune chronic active hepatitis, primary  
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
 CC hypersensitivity, urticaria, an immunologic disease of the lung,  
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
 CC pneumonitis, a transplantation associated disease, graft rejection or  
 CC graft-versus-host disease. The present sequence represents a PRO protein  
 CC of the invention.

XX Sequence 481 AA;

Query Match 96.1%; Score 2315.5; DB 8; Length 481;  
 Best Local Similarity 96.2%; Pred. No. 3.5e-212;  
 Matches 432; Conservative 2; Mismatches 4; Indels 11; Gaps 2;

Qy 8 SH-SVEDTH-----PSHYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRD 56  
 Db 33 SHSSTQENRCKPHQFKESDCHYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRD 92  
 Qy 57 IESGYRIHCSKSVRRSRSPKRNHCSHOSRSKSHRRKRSRSTIEDDEGHLCQSG 116  
 Db 93 IESGYRIHCSKSVRRSRSPKRNHCSHOSRSKSHRRKRSRSTIEDDEGHLCQSG 152  
 Qy 117 DVLARVEIVDTLGEAGFKVVEICDHGMGMHVAVKIVKNGRYREARSEIQVLEHLN 176  
 Db 153 DVLARVEIVDTLGEAGFKVVEICDHGMGMHVAVKIVKNGRYREARSEIQVLEHLN 212  
 Qy 177 STDPSNFRVCVQMLEWDFHGHGVCIVFELLGLSYDFIKENSFLPFQIDHIRMAYQICQ 236  
 Db 213 STDPSNFRVCVQMLEWDFHGHGVCIVFELLGLSYDFIKENSFLPFQIDHIRMAYQICQ 272  
 Qy 237 SINFLHNKLTHDLPENILFVKSDDYVVKVNSKMKRDETLKNTDIKVVDFFGSATYDDE 296  
 Db 273 SINFLHNKLTHDLPENILFVKSDDYVVKVNSKMKRDETLKNTDIKVVDFFGSATYDDE 332  
 Qy 297 HSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYILGFTVFOHDSKEHLAMMERI 356  
 Db 333 HSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILIEYILGFTVFOHDSKEHLAMMERI 392  
 Qy 357 LGPIPOHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLV 416  
 Db 393 LGPIPOHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLV 452  
 Qy 417 RRMLEYDPTORITLDEALQHPFFDLKKK 445  
 Db 453 RRMLEYDPTORITLDEALQHPFFDLKKK 481

RESULT 5  
 AAB69705  
 ID AAB69705 standard; protein; 481 AA.

XX AAB69705;

XX 14-AUG-2001 (first entry)

XX Human cell cycle regulating protein 53.

XX Human; cell cycle regulating protein 53; cancer; blood disease; HIV;  
 KW immunological disease; inflammation.

XX Homo sapiens.

XX WO200130833-A1.

XX

PD 03-MAY-2001.  
 XX 16-OCT-2000; 2000WO-CN000328.  
 XX 22-OCT-1999; 95CN-00119816.  
 XX (SHAN-) SHANGHAI BIO ROAD GENE DEV LTD.  
 XX Mao Y, Xie Y;  
 XX WPI; 2001-300480/31.  
 XX N-PSDB; AAF89402.  
 XX  
 XX New cell cycle-regulating protein 53 and its polynucleotide, applicable  
 PT in diagnosis and treatment of malignant tumor, hemopathy, human  
 PT immunodeficiency virus infection, immunological diseases and various  
 PT inflammation.  
 XX  
 XX Claim 1; Page 23-24; 30pp; Chinese.  
 PS  
 XX The present invention provides the protein and coding sequences for human  
 XX cell cycle regulating protein 53. The sequences can be used in the  
 CC diagnosis and treatment of malignant tumours, haemopathy, human  
 CC immunodeficiency virus (HIV) infection, immunological diseases and  
 CC various types of inflammation. The present sequence is the cell cycle  
 CC regulating protein 53  
 XX  
 XX Sequence 481 AA;  
 SQ  
 Query Match 95.2%; Score 2294.5; DB 4; Length 481;  
 Best Local Similarity 95.5%; Pred. No. 3.5e-210;  
 Matches 429; Conservative 2; Mismatches 7; Indels 11; Gaps 2;  
 Qy 8 SH-SVEEDTH-----PSHYLEARSLSNERDYRDRYVDEYRNDYCEGVPRHHRD 56  
 Db |||||:|:|||||  
 33 SHSTQENRHCKPHQPFKESDCHYLEARSLNERDYRDRYVDEYRNDYCEGVPRHHRG 92  
 Qy 57 IESGYRHCSSKSVRSRSPKRNHRHCSHQSRKSHRKRKRSIEDDEGHLCQSG 116  
 Db |||||  
 93 IESGYRHCSSKSVRSRSPKRNHRHCSHQSRKSHRKRKRSIEDDEGHLCQSG 152  
 Qy 117 DVLARVEIVDTLGEAGKGVKVEICDHGMDGMHVAVKIKVNGRYREARSEIQVLEHLN 176  
 Db |||||  
 153 DVLARVEIVDTLGEAGKGVKVEICDHGMDGMHVAVKIKVNGRYREARSEIQVLEHLN 212  
 Qy 177 STDNSVRCVOMLEWFDHGHVCIVFELLGLSTYDFIKENSLFPQIDHIRQWAYQICQ 236  
 Db |||||  
 213 STDNSVRCVOMLEWFDHGHVCIVFELLGLSTYDFIKENSLFPQIDHIRQWAYQICQ 272  
 Qy 237 SINFLHNKLTHTDKPENILFVKSDYVYKYNKMKRDERLTAKNTDIKVPDFGSATYDDE 296  
 Db |||||  
 273 SINFLHNKLTHTDKPENILFVKSDYVYKYNKMKRDERLTAKNTDIKVPDFGSATYDDE 332  
 Qy 297 HHSTLVSTRHVRAPVILALQWSQPCDWSVSGICILIEYLGFTVFQTHDSKEHLAMMERI 356  
 Db |||||  
 333 HHSTLVSTRHVRAPVILALQWSQPCDWSVSGICILIEYLGFTVFQTHDSKEHLAMMERI 392  
 Qy 357 LGPIPOHMIQTRKRYFHHNOLDWDEHSSAGRYVRCKPLKBFMLCHDEHEKLFDLV 416  
 Db |||||  
 393 LGPIPOHMIQTRKRYFHHNOLDWDEHSSAGRYVRCKPLKBFMLCHDEHEKLFDLV 452  
 Qy 417 RRMLEYDPTQRTITLDEALQHPFFDLKKK 445  
 Db |||||  
 453 RRMLEYDPTQRTITLDEALQHPFFDLKKK 481  
 RESULT 6  
 ADI26141  
 ID ADI26141 standard; protein; 481 AA.  
 XX  
 AC ADI26141;  
 XX  
 DT 22-APR-2004 (first entry)

XX Human protein that promotes STAT6 activation #53.  
 DE human; signal transducer and activator of transcription 6; STAT6;  
 KW immunogen; STAT6 activation; allergy; inflammation; autoimmune disease;  
 KW diabetes; hyperlipidaemia; infection; cancer; Th1 hyperactive disease;  
 KW rheumatoid arthritis; osteoarthritis; systemic lupus erythematosus;  
 KW sepsis; asthma; allergic rhinitis; ischaemic heart disease;  
 KW subarachnoid haemorrhage; viral hepatitis; AIDS.  
 OS Homo sapiens.  
 XX  
 XX WO2003104277-A2.  
 XX 18-DEC-2003.  
 XX 05-JUN-2003; 2003WO-JP007123.  
 XX 05-JUN-2002; 2002JP-00164257.  
 XX 06-JUN-2002; 2002US-0385912P.  
 XX 26-DEC-2002; 2002JP-00377326.  
 XX 27-DEC-2002; 2002US-0436467P.  
 XX 15-MAY-2003; 2003JP-00137505.  
 XX 16-MAY-2003; 2003US-0470836P.  
 XX  
 XX (ASAH ) ASAH KASEI KK.  
 XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;  
 XX WPI; 2004-122214/12.  
 XX N-PSDB; ADI26140.  
 XX  
 XX New signal transducer and activator of transcription 6 activation  
 PT promoting purified protein, for diagnosing and treating disease  
 PT associated with activation/inhibition of transcription factor e.g.  
 PT diabetes and cancer.  
 XX  
 XX Claim 1; SEQ ID NO 106; 1368pp; English.  
 XX  
 XX The invention relates to a purified protein promoting signal transducer  
 CC and activator of transcription 6 activation (STAT6). The protein is  
 CC useful for the producing an antibody, which involves administering the  
 CC protein or its epitope-bearing fragments to a non-human animal as an  
 CC antigen. The nucleic acid is useful for diagnosing a disease or  
 CC susceptibility to a disease related to expression or activity of the  
 CC protein. A transformant expressing the protein is useful for screening  
 CC compounds which inhibit or promote STAT6 activation. A transformant  
 CC expressing the protein is useful for producing a pharmaceutical  
 CC composition. Compositions, antibodies and antisense molecules are useful  
 CC for the treating a disease associated with STAT6 activation such as  
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,  
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful  
 CC for treating disease associated with STAT6 activation and/or prevention  
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid  
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,  
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,  
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6  
 CC activity. The protein or nucleic acid is effectively useful for screening  
 CC compounds for treating and preventing disease associated with excessive  
 CC activation or inhibition of STAT6. The present sequence represents the  
 CC amino acid sequence of a human protein which promotes STAT6 activation.  
 XX  
 XX Sequence 481 AA;  
 SQ  
 Query Match 93.8%; Score 2261.5; DB 8; Length 481;  
 Best Local Similarity 93.5%; Pred. No. 5.1e-207;  
 Matches 420; Conservative 8; Mismatches 10; Indels 11; Gaps 2;  
 Qy 8 SH-SVEEDTH-----PSHYLEARSLSNERDYRDRYVDEYRNDYCEGVPRHHRD 56  
 Db |||||:|:|||||  
 33 SHSTQENRHCKPHQPFKESDCHYLEARSLNERDYRDRYVDEYRNDYCEGVPRHHRD 92  
 Qy 57 IESGYRHCSSKSVRSRSPKRNHRHCSHQSRKSHRKRKRSIEDDEGHLCQSG 116

Db 93 VESTYRIHCSKSSVRSRRSPKRNRPKASHQSHSKSHRRKRSIEDDEGHLCQSG 152  
 Qy 117 DVLARARVEIVDTLGEAGFGKVEICIDHGMCMHVAVKIVKNGRYRAARSEIOVLEHLN 176  
 Db 153 DVLARARVEIVDTLGEAGFGKVEICIDHGMCMHVAVKIVKNGRYRAARSEIOVLEHLN 212  
 Qy 177 STDNSVPRVCQMLEWFDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRMAYOIQ 236  
 Db 213 STDNSVPRVCQMLEWFDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRMAYOIQ 272  
 Qy 237 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDERTLKNVDIFGSGATYDDE 296  
 Db 273 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDERTLKNVDIFGSGATYDDE 332  
 Qy 297 HSTLVSSTRHYRAPEVILALGWSQPCDWSIGCILIEYVLGFTVFQTHDSKEHLAMMERI 356  
 Db 333 HSTLVSSTRHYRAPEVILALGWSQPCDWSIGCILIEYVLGFTVFQTHDSKEHLAMMERI 392  
 Qy 357 LGPIPOHMIQTRKRYFHHNQDWDHSSAGRYVRRCKPLKEFMLECHDEEHEKLFDLV 416  
 Db 393 LGPIPAHMIQTRKRYFHHNQDWDHSSAGRYVRRCKPLKEFMLECHDEEHEKLFDLV 452  
 Qy 417 RRMLEYDPTQITLDEALQHPFFDLKKK 445  
 Db 453 RRMLEYDPTQITLDEALQHPFFDLKKK 481

RESULT 7  
 ADO08029  
 ID ADO08029 standard; protein; 481 AA.  
 AC ADO08029;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Mouse polypeptide #36.  
 XX  
 KW Mouse; fat cell number; fat cell size; obesity; diabetes; anorectic;  
 XX  
 OS Mus sp.  
 XX  
 PN US2004071700-A1.  
 XX  
 PD 15-APR-2004.  
 XX  
 PF 09-OCT-2002; 2002US-00267502.  
 XX  
 PR 09-OCT-2002; 2002US-00267502.  
 XX  
 PA (LIFE-) LIFE SCI DEV CORP.  
 XX  
 PI Kim J, Galant R;  
 XX  
 DR WPI; 2004-328526/30.  
 XX  
 DR N-PSDB; ADO07812.  
 XX

Identifying compounds that influence fat cell number or size for treating or preventing obesity or diabetes by exposing the cell to the agent and identifying fat cell number or size relative to cells not exposed to the agent.  
 Claim 14; SEQ ID NO 355; 275pp; English.  
 The invention relates to a method of identifying compounds that influence fat cell number or size comprising providing a cell that expresses a gene and an agent, exposing the cell to the agent and identifying fat cell number or size relative to cells not exposed to the agent. The method also comprises providing an expression vector and an agent, exposing the vector to the agent, detecting a change in expression of the gene relative to expression of the gene in an expression vector not exposed to the agent, treating a subject with the agent and identifying fat cell

CC number or size in the subject. The agent comprises an antisense CC oligonucleotide. The subject comprises a mammal, preferably a human. The CC method also comprises providing a polypeptide and an agent, exposing the CC polypeptide to the agent, detecting binding of the agent to the CC polypeptide or a change in an activity of the polypeptide, treating a CC subject with the agent and identifying fat cell number or size in the CC subject. The agent comprises an antibody. A method of regulating fat cell CC number or size comprises providing a subject containing fat cells and an CC agent that changes the expression of a gene, and treating the subject CC with the agent under conditions so that fat cell size or number in the CC subject is altered. The method is useful for identifying compounds that CC influence fat cell number or size, for preparing a composition for CC treating or preventing obesity or diabetes. This sequence represents a CC mouse polypeptide used in the scope of the invention.

Qy 8 SH-SVEEDTH-----PSHYLEARSLNERDYDRRYVDEYRNDYCEGYVPRHYHRD 56  
 Db 33 SHSSTGRENHCKEHHQFKSDCHYLEARCLNERDYDRRYVDEYRNDYCEGYVPRHYHRD 92  
 Qy 57 IESGYRIHCSKSSVRSRRSPKRNRPKASHQSHSKSHRRKRSIEDDEGHLCQSG 116  
 Db 93 VESTYRIHCSKSSVRSRRSPKRNRPKASHQSHSKSHRRKRSIEDDEGHLCQSG 152  
 Qy 117 DVLARARVEIVDTLGEAGFGKVEICIDHGMCMHVAVKIVKNGRYRAARSEIOVLEHLN 176  
 Db 153 DVLARARVEIVDTLGEAGFGKVEICIDHGMCMHVAVKIVKNGRYRAARSEIOVLEHLN 212  
 Qy 177 STDNSVPRVCQMLEWFDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRMAYOIQ 236  
 Db 213 STDNSVPRVCQMLEWFDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRMAYOIQ 272  
 Qy 237 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDERTLKNVDIFGSGATYDDE 296  
 Db 273 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDERTLKNVDIFGSGATYDDE 332  
 Qy 297 HSTLVSSTRHYRAPEVILALGWSQPCDWSIGCILIEYVLGFTVFQTHDSKEHLAMMERI 356  
 Db 333 HSTLVSSTRHYRAPEVILALGWSQPCDWSIGCILIEYVLGFTVFQTHDSKEHLAMMERI 392  
 Qy 357 LGPIPOHMIQTRKRYFHHNQDWDHSSAGRYVRRCKPLKEFMLECHDEEHEKLFDLV 416  
 Db 393 LGPIPAHMIQTRKRYFHHNQDWDHSSAGRYVRRCKPLKEFMLECHDEEHEKLFDLV 452  
 Qy 417 RRMLEYDPTQITLDEALQHPFFDLKKK 445  
 Db 453 RRMLEYDPTQITLDEALQHPFFDLKKK 481

RESULT 8  
 AAW49914  
 ID AAW49914 standard; protein; 479 AA.  
 XX  
 AC AAW49914;  
 XX  
 DT 20-JUL-1998 (first entry)  
 XX  
 DE Mouse CLK serine/threonine kinase mCLK4.  
 XX  
 KW mCLK4; CLK; serine/threonine kinase; protein kinase; LAMMER kinase;  
 XX signal transduction; cancer; contraceptive; mouse; therapy; diagnosis.  
 XX  
 OS Mus musculus.  
 XX  
 PH Key Location/Qualifiers  
 FT Domain 29..44  
 FT Domain /note= "nuclear localisation domain"  
 FT Domain 159..476



| FT  | Peptide   | /note= "catalytic domain" |
|-----|---|---------------------------|
| FT  | 384..389  |                           |
| FT  | /note= "LAMMER motif"   |                           |
| XX  | WO9748723-A2..  |                           |
| XX  | 24-DEC-1997.  |                           |
| XX  | 17-JUN-1997;  | 97WO-IB000946.            |
| XX  | 17-JUN-1996;  | 96US-0019629P.            |
| PR  | 09-AUG-1996;  | 96US-0023485P.            |
| PR  | 13-NOV-1996;  | 96US-0030860P.            |
| PR  | 15-NOV-1996;  | 96US-0030964P.            |
| PR  | 19-DEC-1996;  | 96US-0034286P.            |
| XX  | (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.   |                           |
| PA  | Ullrich A, Kharitonov AI, Aoki N, Wang HY, Chen Z, Nayler O;  |                           |
| PI  | Kim YW;   |                           |
| PI  | WPI; 1998-120302/11.  |                           |
| XX  | New phosphatase and kinase enzyme(s) - useful in the diagnosis and treatment of signal transduction disorders.  |                           |
| XX  | Claim 11; Fig 4; 138pp; English.  |                           |
| XX  | This polypeptide comprises novel mouse CLK serine/threonine kinase mCLK4, from the CLK serine/threonine kinase family of proteins that regulate RNA splicing in cells. mCLK4 cDNA was cloned from a mouse embryo 11.5 p.c. 12AP CDNA library. The invention discloses the discovery of novel protein kinases mCLK2 (see AAW49912), mCLK3 and mCLK4 (see AAW49914) of mol.wt. 59.9, 58.5 and 57.2 kDa, respectively, as well as other novel proteins (see AAW49906-10) involved in cellular signal transduction, and provides vectors, host cells, purified recombinant proteins, methods for identifying compounds that activate or inhibit the novel proteins, as well as methods for the diagnosis and treatment of diseases associated with the novel proteins. Overexpression of CLK serine/threonine kinases has been implicated in certain types of cancer. Compounds that inhibit their catalytic activity or disrupt their interactions with natural binding partners may act as anti-cancer therapeutics. mCLK related molecules and compounds may also be useful as male contraceptives |                           |
| XX  | Sequence 479 AA;  |                           |
| SQL | Query Match 90.4%; Score 2177.5; DB 2; Length 479;  |                           |
|     | Best Local Similarity 91.1%; Pred. No. 5.4e-199;  |                           |
|     | Matches 409; Conservative 9; Mismatches 18; Indels 13; Gaps 4;  |                           |
| QY  | 8 SH-SVEEDTH-----PSHYLEARSINERDYDRRYVDEVYNDYCEGVPRHYRD 56   |                           |
| DB  | 33 SHSTQENRCKPHQFQKSDCHYLEARCLNERYDR-RYDEVYNDYCEGVPRHYRD 91   |                           |
| QY  | 57 IESGYRIHCSKSSVRSRRSPKRRNRHSCSHQSRKSRSHRRKRGRSIEDDEEGLICQSG 116   |                           |
| DB  | 92 VESTYRIHCSKSSVRSRRSPKRRNRPCASHQSHSHRRKRGRSIEDDEEGLICQSG 151  |                           |
| QY  | 117 DVLRLARYEIVDTLGEAGFGKVEICIDHGMQGHVAVKIVKNGVRYREARSEIQVLEHLN 176   |                           |
| DB  | 152 DVLRLARYEIVDTLGEAGFGKVEICIDHGMQGLHVAVKIVKNGVGYREARSEIQVLEHLN 211  |                           |
| QY  | 177 STDPNSVFRVCQMLEWFDHGHGVCIVFELLGLSTYDFTIKENSFLPQIDHIRQWAYCIQC 236  |                           |
| DB  | 212 STDPNSVFRVCQMLEWFDHGHGVCIVFELLGLSTYDFTIKENSFLPQIDHIRQWAYCIQC 271  |                           |
| QY  | 237 SINFLHNKLTHTDLKPNILFVKSDYVYKYNKMKRDERLTAKNTDIKVDVFGSATYDDE 296  |                           |
| DB  | 272 SINFLHNKLTHTDLKPNILFVKSDYVYK-NPSMKRDERLTAKNTIKVDVFGSATYDDE 330  |                           |
| QY  | 297 HHSTLVSTRYRAPVETIALGWSQPCDWSVTCGILLIEYLGFTVFTQTHDSKEHLAMMERI 356  |                           |
| DB  | 331 HHSTLVSTRYRAPVETIALGWSQPCDWSVTCGILLIEYLGFTVFTQTHDSKEHLAMMERI 390  |                           |

QY 8 SH-SVEEDTH-----PSHYLEARSINERDYYVDEYRNDYCEGVVPRHYRD 56  
 DB 81 SHSSTQENRHCKPHQPKESDCHYLEARSINERDYYVDEYRNDYCEGVVPRHYRD 140  
 QY 57 IESGYRIHCSKSSVRSRRSPKRNHRHCSHQSRKSHRRKRSRSTEDDEGHLICQSG 116  
 DB 141 IESGYRIHCSKSSVRSRRSPKRNHRHCSHQSRKSHRRKRSRSTEDDEGHLICQSG 176  
 QY 117 DVLBARVEIVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGVRYREARSEIOVLEHLN 176  
 DB 177 -----XEVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGVRYREARSEIOVLEHLN 230  
 QY 177 STDNSVFRVCQVLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPQIDHIRQAYQICQ 236  
 DB 231 STDNSVFRVCQVLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPQIDHIRQAYQICQ 290  
 QY 237 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDEKTLKNTDIKVVDFGSATYDDE 296  
 DB 291 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDEKTLKNTDIKVVDFGSATYDDE 350  
 QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 356  
 DB 351 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 410  
 QY 357 LGPIPOHMIQKTRKRYFHHNQLDWEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLPDLV 416  
 DB 411 LGPIPOHMIQKTRKRYFHHNQLDWEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLPDLV 470  
 QY 417 RRMLEYDPTQRTILDEALQHPFFDLKKK 445  
 DB 471 RRMLEYDPTQRTILDEALQHPFFDLKKK 499

RESULT 10  
 ID ADI29255 standard; protein; 499 AA.  
 AC ADI29255;  
 XX  
 DT 22-APR-2004 (first entry)  
 DE Human MARK3-associated protein #40.  
 XX  
 KW Human; antisense gene therapy; MARK3;  
 KW MAP/microtubule affinity-regulating kinase 3; cancer;  
 KW Alzheimer's disease; neurodegenerative disorder;  
 KW hyperproliferative disorder; cytostatic.  
 OS Homo sapiens.  
 XX  
 PN US2003232771-A1.  
 XX  
 PD 18-DEC-2003.  
 XX  
 PF 17-JUN-2002; 2002US-00174319.  
 XX  
 PR 17-JUN-2002; 2002US-00174319.  
 XX  
 PA (ISIS-) ISIS PHARM INC.  
 XX  
 PI Ward DT, Freier SM, Dobie KW;  
 XX  
 DR WPI; 2004-052188/05.  
 DR N-ESDB; ADI29373.  
 XX  
 XX New antisense compound targeted to a nucleic acid molecule encoding  
 PT microtubule-affinity-regulating kinases (MARK3); useful for modulating  
 PT expression of MARK3 or for treating cancer or Alzheimer's disease.  
 XX  
 PS Disclosure; SEQ ID NO 175; 233pp; English.  
 XX  
 CC The invention relates to a compound comprising a sequence comprising 8-80

CC base pairs (bp) targeted to a nucleic acid encoding MARK3  
 CC (MAP/microtubule affinity-regulating kinase 3), that specifically  
 CC hybridizes with the nucleic acid encoding MARK3 and inhibits expression  
 CC of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a  
 CC composition comprising the compound and a carrier or diluent, inhibiting or  
 CC the expression of MARK3 in cells or tissues, treating an animal having or  
 CC suspected of having a disease or condition associated with MARK3 and  
 CC screening for an antisense compound. The antisense oligonucleotide is  
 CC useful for preparing a composition for treating hyperproliferative  
 CC disorder, particularly cancer and neurodegenerative diseases e.g.  
 CC Alzheimer's disease. The present sequence is a MARK3 associated protein  
 CC included in the figures but not mentioned anywhere else in the  
 CC specification.  
 XX Sequence 499 AA;  
 SQ

Query Match 88.3%; Score 2127.5; DB 8; Length 499;  
 Best Local Similarity 89.3%; Pred. No. 3.5e-194;  
 Matches 401; Conservative 2; Mismatches 5; Indels 41; Gaps 3;

QY 8 SH-SVEEDTH-----PSHYLEARSINERDYYVDEYRNDYCEGVVPRHYRD 56  
 DB 81 SHSSTQENRHCKPHQPKESDCHYLEARSINERDYYVDEYRNDYCEGVVPRHYRD 140  
 QY 57 IESGYRIHCSKSSVRSRRSPKRNHRHCSHQSRKSHRRKRSRSTEDDEGHLICQSG 116  
 DB 141 IESGYRIHCSKSSVRSRRSPKRNHRHCSHQSRKSHRRKRSRSTEDDEGHLICQSG 176  
 QY 117 DVLBARVEIVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGVRYREARSEIOVLEHLN 176  
 DB 177 -----XEVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGVRYREARSEIOVLEHLN 230  
 QY 177 STDNSVFRVCQVLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPQIDHIRQAYQICQ 236  
 DB 231 STDNSVFRVCQVLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPQIDHIRQAYQICQ 290  
 QY 237 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDEKTLKNTDIKVVDFGSATYDDE 296  
 DB 291 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKMKRDEKTLKNTDIKVVDFGSATYDDE 350  
 QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 356  
 DB 351 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 410  
 QY 357 LGPIPOHMIQKTRKRYFHHNQLDWEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLPDLV 416  
 DB 411 LGPIPOHMIQKTRKRYFHHNQLDWEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLPDLV 470  
 QY 417 RRMLEYDPTQRTILDEALQHPFFDLKKK 445  
 DB 471 RRMLEYDPTQRTILDEALQHPFFDLKKK 499

RESULT 11  
 AAY27054  
 ID AAY27054 standard; protein; 451 AA.  
 XX  
 AC AAY27054;  
 XX  
 DT 08-OCT-1999 (first entry)  
 DE Human protein kinase (HPKM)-3 (clone ID 339963).  
 XX  
 DE Human protein kinase molecule; HPKM; human; protein kinase;  
 KW phosphate group; cancer; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 4  
 FT Modified-site 22 /note= "potential phosphorylation site"  
 FT Modified-site 22 /note= "potential phosphorylation site"





| Matches | 352; Conservative | 30; Mismatches  | 45; Indels | 2; Gaps | 2; |
|---------|-------------------|---|------------|---------|----|
| Qy      | 18                | SHYLEARSLNERDYRDRYRYVDEYENDYCEGYVPRHYRHDIESGYRIHCSSKSVSRSSSP  | 77         |         |    |
| Db      | 54                | SHYLESRSINEKDYHSRRYIDEYRNDYTQGCPCGHRQRDHESRYQNHSSKSSGSRSSSY   | 113        |         |    |
| Qy      | 78                | KRK-RNPHCSH-QSRKSHRRKRSIEDDEEGHLICQSGDVLARYEIVDTLGEAGF        | 135        |         |    |
| Db      | 114               | KSKRHHSSTSHRRSGKSHRRKTRSEVDEEGHLICQSGDVLARSYEIVDTLGEAGF       | 173        |         |    |
| Qy      | 136               | KWCECIDHGMGMHVAVKIKVNGRYREAAARSEIQVLEHLNSTDPNSVFCQVMLEWFDH    | 195        |         |    |
| Db      | 174               | KWCECIDHKGAGRHVAVKIKVNDRYCEAAARSEIQVLEHLNTTDPNSTFCVQMLEWFH    | 233        |         |    |
| Qy      | 196               | HGHVCIYPELLGLSTYDFIKENSLFPQIDHIROMAYQICOSINFLHNKLTHTDLPEN     | 255        |         |    |
| Db      | 234               | HGHICVYPELLGLSTYDFIKENGLFPLFDHIRMAYQICKSVNFLHSNKLTHTDLPEN     | 293        |         |    |
| Qy      | 256               | ILFVKSDYVWKYNSMKRDERLTAKNTDIKVWDFGSATYDDEHHSGLVSTRHYRAPEVILA  | 315        |         |    |
| Db      | 294               | ILFQSDYTEAYNPKIKRDERTLNPDIKVWDFGSATYDDEHHSGLVSTRHYRAPEVILA    | 353        |         |    |
| Qy      | 316               | LGWSQPCDVMSIGCILIEYILGYTFVPTHDSKEHLAMMERILGPIPOHMIQKTRKRYFH   | 375        |         |    |
| Db      | 354               | LGWSQPCDVMSIGCILIEYILGYTFVPTHDSKEHLAMMERILGLPLPKHMIQKTRKRYFH  | 413        |         |    |
| Qy      | 376               | HNQLDWDHSSAGRYVRRRCPLKPEFMLCHDBEHEKFLDILVRRMLEYDPTQIRITLDRALQ | 435        |         |    |
| Db      | 414               | HDRLDWDHSSAGRYVSRACKPLKBFMLSQDVEHERFLDILIQKLEFYDPAKRIITLRALK  | 473        |         |    |
| Qy      | 436               | HPFPDLKK  | 444        |         |    |
| Db      | 474               | HPFPDLKK  | 482        |         |    |

|           |   |
|-----------|---|
| RESULT 14 |   |
| ADH58849  |   |
| ID        | ADH58849 standard; protein; 484 AA.                                     |
| XX        |   |
| AC        | ADH58849;   |
| XX        |   |
| DT        | 25-MAR-2004 (first entry)   |
| XX        |   |
| DE        | Human CDC-like kinase-1 protein #1.                                     |
| XX        |   |
| KW        | antitense oligonucleotide; CDC-like kinase 1; cancer;                   |
| KW        | autoimmune disorder; infection; inflammation; tumour formation; human.  |
| XX        |   |
| OS        | Homo sapiens.   |
| XX        |   |
| PN        | US2003219895-A1.  |
| XX        |   |
| PD        | 27-NOV-2003.  |
| XX        |   |
| PF        | 22-MAY-2002; 2002US-00154708.   |
| XX        |   |
| PR        | 22-MAY-2002; 2002US-00154708.   |
| XX        |   |
| PA        | (ISIS-) ISIS PHARM INC.   |
| XX        |   |
| PI        | Watt AT;  |
| XX        |   |
| DR        | WPI; 2004-051714/05.  |
| DR        | N-PSDB; ADH58708.   |
| XX        |   |
| PT        | New antisense oligonucleotides targeted to nucleic acid molecules...    |
| PT        | encoding CDC-like kinase 1, useful for treating diseases or conditions  |
| PT        | associated with expression of CDC-like kinase 1, e.g. cancers or        |
| PT        | autoimmune disorders.   |
| XX        |   |
| PS        | Example 13; Page 33-35; 64pp; English.                                  |
| XX        |   |
| CC        | The invention comprises antisense oligonucleotides that are targeted to |

[illegible]

KW HIV propagation; human.  
XX Homo sapiens.  
XX WO2004065577-A2.  
XX 05-AUG-2004.  
XX 13-JAN-2004; 2004WO-US000798.  
XX 14-JAN-2003; 2003US-0440068P.  
XX 12-MAY-2003; 2003US-0469757P.  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX Nadler SG, Neubauer MG, Feder JN, Carman J;  
XX WPI; 2004-562168/54.  
XX N-ESDB; ADR14000.  
XX New isolated polynucleotides and polypeptides associated with NF-kappaB  
XX pathway, useful for diagnosing, treating, or preventing disorders or  
XX diseases associated with NF-kappaB pathway.  
XX Claim 6; SEQ ID NO 2; 237pp; English.  
XX This invention relates to the novel association of protein sequences (and  
XX the genes which encode them) to the NF-kappaB pathway. The invention may  
XX be useful for the production of compounds with an antiinflammatory,  
XX cytosolic, hepatotropic, virucide, antiarthritic, antirheumatic,  
XX gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,  
XX immunomodulator, cerebroprotective, vasotropic, immunosuppressive or  
XX vulnary activity or for gene therapy. The proteins and nucleotides are  
XX useful for diagnosing, preventing, treating, or ameliorating conditions  
XX or diseases associated with the NF-kappaB pathway. The condition is an  
XX immune disorder, an inflammatory disorder, an inflammatory disorder  
XX related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,  
XX hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM  
XX syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic  
XX ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,  
XX hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell  
XX survival, evasion of immune responses, rheumatoid arthritis, inflammatory  
XX bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
XX syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
XX immune activity, disorders related to aberrant acute phase responses,  
XX hypercongenital conditions, birth defects, necrotic lesions, wounds,  
XX organ transplant rejection, conditions related to organ transplant  
XX rejection, disorders related to aberrant signal transduction,  
XX proliferating disorders, cancers and HIV propagation in cells infected  
XX with other viruses. The present sequence is that of a human protein which  
XX is subject to the novel association with the NF-kappaB pathway of the  
XX invention. Note: This sequence does not appear in the specification but  
XX was obtained by the indexer from Genbank.  
XX SQ Sequence 484 AA;  
Query Match 78.3%; Score 1887; DB 8; Length 484;  
Best Local Similarity 82.1%; Pred. No. 3.3e-171;  
Matches 352; Conservative 30; Mismatches 45; Indels 2; Gaps 2;  
Qy 18 SHYLEARSLNRDRYRDYDEYNDYCEGVPRHYHRDIESGYRHCSKSVSRSSP 77  
Db 54 SHYLESRSINEKDYHSRRYIDEYNDYTQCGEPGHRQDRHESRYQNHSSKSGRSY 113  
Qy 78 KRK-RNRHCSH-QSRSKSHRKRKRSSTDEEGHLICQSGDVLRYEIVDTLGEAGF 135  
Db 114 KSKRIHHTSHRRSHGSHKRRKRTSRVEDDEGHLICQSGDVLRYEIVDTLGEAGF 173  
Qy 136 KVECEIDHGMGMVAVKIVKNGRYEARSEIQVLEHLNSTDPNSVRCVQMLEWFDH 195  
Db 174 KVECEIDHKGRRHVAVKIVKNDRYCEARSEIQVLEHLNTDPNSTFRVCVQMLEWFEH 233  
Qy 196 HGHVCIIVFELLGLSTYDFIKENSFLPFDIHIQWMAQYQICQSNFLHKNKLTHTDLKPEN 255

Db 234 HGHICIVFELLGLSTYDFIKENGFLPFDLHIKMAQYQICKSYNFLHSNKLTHTLKPEN 293  
Qy 256 ILFVKSDYVVKYNSKMKRDERTLKNDDIKVDFGSATYDDEHSTLVSTRHYRAPEVILA 315  
Db 294 ILFQSDYTEAYNPKIKRDERTLINDDIKVDFGSATYDDEHSTLVSTRHYRAPEVILA 353  
Qy 316 LGWSQPCDVMSIGCILIEYYLGFVFTQTHDSKEHLAMMERILGPIPOHMIQTKRKRYFH 375  
Db 354 LGWSQPCDVMSIGCILIEYYLGFVFTQTHDSKEHLAMMERILGPIPOHMIQTKRKRYFH 413  
Qy 376 HNOIDWDEHSSAGRYVRRCKPLKEPMLCHDEHEKLFDLVRRMLEYDPTQRTILDEALQ 435  
Db 414 HDRLDWDEHSSAGRYVSRACKPLKEPMLSQDVEHERLFDLIQWLEYDPAKRITLREALK 473  
Qy 436 HPFFDLLKK 444  
Db 474 HPFFDLLKK 482

Search completed: March 13, 2005, 00:05:04  
Job time : 170 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 14:32:04 ; Search time 1259.64 Seconds  
(without alignments)  
10662.144 Million cell updates/sec

Title: US-10-801-671-1\_COPY\_72\_2327

Perfect score: 2256

Sequence: 1 gacactcatccagtcattt.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Match | Length | DB | ID                  | Description       |
|------------|--------|-------|--------|----|---------------------|-------------------|
| 1          | 2256   | 100.0 | 2354   | 9  | US-09-810-671-1     | Sequence 1, Appli |
| 2          | 2256   | 100.0 | 2354   | 13 | US-10-109-854-1     | Sequence 1, Appli |
| 3          | 2256   | 100.0 | 2354   | 15 | US-10-339-656-1     | Sequence 1, Appli |
| 4          | 2256   | 100.0 | 2354   | 18 | US-10-801-671-1     | Sequence 1, Appli |
| 5          | 1908.4 | 84.6  | 4035   | 17 | US-10-425-114-26852 | Sequence 26852, A |
| 6          | 1286   | 57.0  | 1446   | 17 | US-10-267-502-135   | Sequence 135, App |
| 7          | 1178.2 | 52.2  | 1456   | 17 | US-10-641-643-699   | Sequence 699, App |
| 8          | 1097.4 | 48.6  | 1549   | 9  | US-09-905-999-26    | Sequence 26, Appl |
| 9          | 1097.4 | 48.6  | 1549   | 18 | US-10-825-177-26    | Sequence 26, Appl |
| 10         | 1097.2 | 48.6  | 1446   | 17 | US-10-267-502-138   | Sequence 138, App |
| 11         | 1097   | 48.6  | 21234  | 9  | US-09-810-671-3     | Sequence 3, Appli |

|    |       |      |       |    |                     |                   |
|----|-------|------|-------|----|---------------------|-------------------|
| 12 | 1097  | 48.6 | 21234 | 13 | US-10-109-854-3     | Sequence 3, Appli |
| 13 | 1097  | 48.6 | 21234 | 15 | US-10-339-656-3     | Sequence 3, Appli |
| 14 | 1097  | 48.6 | 21234 | 18 | US-10-801-671-3     | Sequence 3, Appli |
| 15 | 977.4 | 43.3 | 3040  | 17 | US-10-425-114-26212 | Sequence 26212, A |
| 16 | 813.2 | 36.0 | 1455  | 17 | US-10-267-502-136   | Sequence 136, App |
| 17 | 811.6 | 36.0 | 1834  | 17 | US-10-154-708-3     | Sequence 3, Appli |
| 18 | 811.6 | 36.0 | 1834  | 18 | US-10-755-889-1     | Sequence 1, Appli |
| 19 | 768.8 | 34.1 | 1452  | 17 | US-10-267-502-139   | Sequence 139, App |
| 20 | 768.4 | 34.1 | 906   | 17 | US-10-267-502-140   | Sequence 140, App |
| 21 | 644.4 | 28.6 | 1743  | 17 | US-10-154-708-10    | Sequence 10, Appl |
| 22 | 640.6 | 28.4 | 2254  | 9  | US-09-919-039-238   | Sequence 238, App |
| 23 | 640   | 28.4 | 2516  | 9  | US-09-925-298-121   | Sequence 121, App |
| 24 | 640   | 28.4 | 2516  | 14 | US-10-102-806-121   | Sequence 121, App |
| 25 | 499.6 | 22.1 | 1500  | 10 | US-09-790-852-2     | Sequence 2, Appli |
| 26 | 499.6 | 22.1 | 1973  | 9  | US-09-962-436-366   | Sequence 266, App |
| 27 | 499.6 | 22.1 | 1973  | 9  | US-09-880-107-2190  | Sequence 2190, Ap |
| 28 | 499.6 | 22.1 | 1973  | 18 | US-10-737-450-131   | Sequence 131, App |
| 29 | 481.2 | 21.3 | 1538  | 9  | US-09-905-999-22    | Sequence 22, Appl |
| 30 | 481.2 | 21.3 | 1538  | 18 | US-10-825-177-22    | Sequence 22, Appl |
| 31 | 464.4 | 20.6 | 1296  | 17 | US-10-182-243-16    | Sequence 16, Appl |
| 32 | 464.4 | 20.6 | 1628  | 17 | US-10-425-114-26266 | Sequence 26266, A |
| 33 | 461.6 | 20.5 | 1026  | 17 | US-10-267-502-133   | Sequence 133, App |
| 34 | 458.6 | 20.3 | 1885  | 17 | US-10-439-703-89    | Sequence 89, Appl |
| 35 | 454.8 | 20.2 | 1787  | 9  | US-09-905-999-24    | Sequence 24, Appl |
| 36 | 454.8 | 20.2 | 1787  | 18 | US-10-825-177-24    | Sequence 24, Appl |
| 37 | 453.2 | 20.1 | 1473  | 17 | US-10-267-502-137   | Sequence 137, App |
| 38 | 432.4 | 19.2 | 1473  | 17 | US-10-267-502-134   | Sequence 134, App |
| 39 | 432.4 | 19.2 | 1762  | 17 | US-10-305-720-1439  | Sequence 1439, Ap |
| 40 | 430.8 | 19.1 | 2505  | 17 | US-10-108-260A-2256 | Sequence 2256, Ap |
| 41 | 429.2 | 19.0 | 2621  | 17 | US-10-104-047-656   | Sequence 656, App |
| 42 | 377.4 | 16.7 | 1181  | 17 | US-10-425-114-16170 | Sequence 16170, A |
| 43 | 369   | 16.4 | 369   | 9  | US-09-796-692-5456  | Sequence 5456, Ap |
| 44 | 369   | 16.4 | 369   | 14 | US-10-040-862-5456  | Sequence 5456, Ap |
| 45 | 369   | 16.4 | 369   | 17 | US-10-057-475B-5456 | Sequence 5456, Ap |

#### ALIGNMENTS

##### RESULT 1

US-09-810-671-1  
; Sequence 1, Application US/09810671  
; Publication No. US20020076783A1  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL000758  
; CURRENT APPLICATION NUMBER: US/09/810,671  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2354  
; TYPE: DNA  
; ORGANISM: Human  
US-09-810-671-1

Query Match 100.0%; Score 2256; DB 9; Length 2354;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |  |     |
|----|-----|--|-----|
| Qy | 1   | GACACTCATCCAGTCATTATTAGAACGAGTCCCTTGAATGACGAGATTCGGAC        | 60  |
| Db | 72  | GACACTCATCCAGTCATTATTAGAACGAGTCCCTTGAATGACGAGATTCGGAC        | 131 |
| Qy | 61  | CGGAGATACGTTGCGAATACAGGAATCACTACTGTGAAGGATATGTTCCCTAGACATTAT | 120 |
| Db | 132 | CGGAGATACGTTGCGAATACAGGAATCACTACTGTGAAGGATATGTTCCCTAGACATTAT | 191 |
| Qy | 121 | CACAGAGACATTGAAAGCGGGTATCGAATCCACTGCAGTAAATCTTTCAGTCGACGAGG  | 180 |

Db 192 CACAGAGACATTTGAAGCGGGTATCGAATCCACATGCACTAAATCTTTCAGTCGCGCAGG 251  
 Qy 181 AGAGCAGTCTTAAAGGAAGCGCAATAGACACTGCTCAAGTCATCAGTCAGTCGGAAG 240  
 Db 252 AGAGCAGTCTTAAAGGAAGCGCAATAGACACTGCTCAAGTCATCAGTCAGTCGGAAG 311  
 Qy 241 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATAGGAGGGTCACCTGATCTGT 300  
 Db 312 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATAGGAGGGTCACCTGATCTGT 371  
 Qy 301 CAAGTGGAGAGCGTTCTAAGAGCAAGATAGAAATCGGTGGACACTTTGGGTGAAGAGCC 360  
 Db 372 CAAGTGGAGAGCGTTCTAAGAGCAAGATAGAAATCGGTGGACACTTTGGGTGAAGAGCC 431  
 Qy 361 TTTGSCAAAGTTGTAGAGTGCATTCATCATGSCATGGATGSCATGATAGCAGTGAAG 420  
 Db 432 TTTGSCAAAGTTGTAGAGTGCATTCATCATGSCATGGATGSCATGATAGCAGTGAAG 491  
 Qy 421 ATCGTAAATAATGTAGGCGGTTACCGTGAAGCAGCTCGTTGAGAAATCCAAAGTATTAGAG 480  
 Db 492 ATCGTAAATAATGTAGGCGGTTACCGTGAAGCAGCTCGTTGAGAAATCCAAAGTATTAGAG 551  
 Qy 481 CACTTAAATAGTACTGATCCCAATAGTGTCTTCGATGTGCCAGATGCTAGATGGTTT 540  
 Db 552 CACTTAAATAGTACTGATCCCAATAGTGTCTTCGATGTGCCAGATGCTAGATGGTTT 611  
 Qy 541 GATCATCATGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600  
 Db 612 GATCATCATGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 671  
 Qy 601 ATTTAAAGAAAAACAGCTTCTGCCATTTCCAAATGACCAATGACCAATGACCAATGACCA 660  
 Db 672 ATTTAAAGAAAAACAGCTTCTGCCATTTCCAAATGACCAATGACCAATGACCAATGACCA 731  
 Qy 661 ATCTGCCAGTCAATAAATTTTACATCAATAAATTTTACATCAATAAATTTTACATCAAT 720  
 Db 732 ATCTGCCAGTCAATAAATTTTACATCAATAAATTTTACATCAATAAATTTTACATCAAT 791  
 Qy 721 GAAATATTTTGTGTCAGTCTGACATGTCAGTCAATATATATTTTAAATGAAACGT 780  
 Db 792 GAAATATTTTGTGTCAGTCTGACATGTCAGTCAATATATATTTTAAATGAAACGT 851  
 Qy 781 GATGAACGCACACTGAAACCAACAGATATCAAGTTGTTGATTTGGAAGTGCACAGTAT 840  
 Db 852 GATGAACGCACACTGAAACCAACAGATATCAAGTTGTTGATTTGGAAGTGCACAGTAT 911  
 Qy 841 GATGATGAACATCAGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900  
 Db 912 GATGATGAACATCAGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 971  
 Qy 901 TTGGCTTTAGGT 960  
 Db 972 TTGGCTTTAGGT 1031  
 Qy 961 TATTACCTTTGTTTCCAGTCTTTCCAGTCTTTCCAGTCTTTCCAGTCTTTCCAGTCTTTCC 1020  
 Db 1032 TATTACCTTTGTTTCCAGTCTTTCCAGTCTTTCCAGTCTTTCCAGTCTTTCCAGTCTTTCC 1091  
 Qy 1021 GAAACGAATATTAGACCCATACCAACACATGATTTAGAAACACAGAAACACAGAAACAC 1080  
 Db 1092 GAAACGAATATTAGACCCATACCAACACATGATTTAGAAACACAGAAACACAGAAACAC 1151  
 Qy 1081 TTTTACCAATAACAGCTAGATTGGATGAACACAGTCTGTGTGTGTGTGTGTGTGTGTGTGT 1140  
 Db 1152 TTTTACCAATAACAGCTAGATTGGATGAACACAGTCTGTGTGTGTGTGTGTGTGTGTGTGT 1211  
 Qy 1141 CGCTGCAAAACCGTTGAAGAAATTTATGCTTTGTCATGATGAAGAACATGAGAAACCTGTTT 1200  
 Db 1212 CGCTGCAAAACCGTTGAAGAAATTTATGCTTTGTCATGATGAAGAACATGAGAAACCTGTTT 1271  
 Qy 1201 GACCTGGTTTCGAGAAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCA 1260  
 Db 1272 GACCTGGTTTCGAGAAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCA 1331

Qy 1261 TTGAGCAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320  
 Db 1332 TTGAGCAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1391  
 Qy 1321 TATATACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380  
 Db 1392 TATATACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1451  
 Qy 1381 TTTTGTAAACATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440  
 Db 1452 TTTTGTAAACATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1511  
 Qy 1441 TAGCATTAATTAACCTTGTAAAGCAAGTATGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1500  
 Db 1512 TAGCATTAATTAACCTTGTAAAGCAAGTATGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1571  
 Qy 1501 AATTTTCT 1560  
 Db 1572 AATTTTCT 1631  
 Qy 1561 ATAAATGTGATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1620  
 Db 1632 ATAAATGTGATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1691  
 Qy 1621 AGTAAAGGAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680  
 Db 1692 AGTAAAGGAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1751  
 Qy 1681 TTTTGAACCTTAACTTTAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740  
 Db 1752 TTTTGAACCTTAACTTTAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1811  
 Qy 1741 CTCTAGATTAAGCAGTACTGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1800  
 Db 1812 CTCTAGATTAAGCAGTACTGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1871  
 Qy 1801 TAAATTTTAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860  
 Db 1872 TAAATTTTAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1931  
 Qy 1861 ATGAAAGGCAATGCAAGTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTT 1920  
 Db 1932 ATGAAAGGCAATGCAAGTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTT 1991  
 Qy 1921 TTTTAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980  
 Db 1992 TTTTAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2051  
 Qy 1981 TTTTGTGAGTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2040  
 Db 2052 TTTTGTGAGTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2111  
 Qy 2041 AAGTCTTTAGAGTCTTAAACATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCAT 2100  
 Db 2112 AAGTCTTTAGAGTCTTAAACATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCAT 2171  
 Qy 2101 ACCTATTAAGATTAAGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160  
 Db 2172 ACCTATTAAGATTAAGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2231  
 Qy 2161 AAAGAGGATTTGATTTGGAAGCAGTTTGGGGAAGAAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
 Db 2232 AAAGAGGATTTGATTTGGAAGCAGTTTGGGGAAGAAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2291  
 Qy 2221 AATTGATTTGTTTACATAAACTTTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2256  
 Db 2292 AATTGATTTGTTTACATAAACTTTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT



```
; Publication No. US200201195481
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-1

Query Match      100.0%; Score 2256; DB 13; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACTCATCCAGTCATATTATTAGAACGAAGTCTTTGAATGAGCGAGATTATCGGGAC 60
DB 72 GACACTCATCCAGTCATATTATTAGAACGAAGTCTTTGAATGAGCGAGATTATCGGGAC 131

QY 61 CGGAGATACGTTGCGAATACAGGAATCACTACTGTGAGGATATGTTCTTAGACATTAT 120
DB 132 CGGAGATACGTTGCGAATACAGGAATCACTACTGTGAGGATATGTTCTTAGACATTAT 191

QY 121 CACAGAGACATTGGAAGCGGGTATCGAATCCACTGCGTAAATCTTCACTCGCGAGCAGG 180
DB 192 CACAGAGACATTGGAAGCGGGTATCGAATCCACTGCGTAAATCTTCACTCGCGAGCAGG 251

QY 181 AGAAGCAGTCTTAAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTTCGAAG 240
DB 252 AGAAGCAGTCTTAAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTTCGAAG 311

QY 241 AGCCACCAAGGAAGATCAGGATGATAGAGATGATAGGAGGTCACCTGATCTGT 300
DB 312 AGCCACCAAGGAAGATCAGGATGATAGAGATGATAGGAGGTCACCTGATCTGT 371

QY 301 CAAAGTGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 360
DB 372 CAAAGTGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 431

QY 361 TTGGCAAAGTTGTAGAGTCATTTGATCATGGCATGGATGGCATGTCATGTAGCAGTGAAA 420
DB 432 TTGGCAAAGTTGTAGAGTCATTTGATCATGGCATGGATGGCATGTCATGTAGCAGTGAAA 491

QY 421 ATCGTAAAGATGATAGGCGGTTACCGTGAAGCAGCTCGTTGAGAAATCCAGTATTAGAG 480
DB 492 ATCGTAAAGATGATAGGCGGTTACCGTGAAGCAGCTCGTTGAGAAATCCAGTATTAGAG 551

QY 481 CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAGATGGTTT 540
DB 552 CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAGATGGTTT 611

QY 541 GATCATCATGGTCATGTTGTATTGTTGTTGAATCTACTGGGACTTTAGTACTTACGATTTTC 600
DB 612 GATCATCATGGTCATGTTGTATTGTTGTTGAATCTACTGGGACTTTAGTACTTACGATTTTC 671

QY 601 ATTAAGAAAACAGCTTTCTGCGATTTCAAAATGACCAATCAGGCAGATGCGGTATCAG 660
DB 672 ATTAAGAAAACAGCTTTCTGCGATTTCAAAATGACCAATCAGGCAGATGCGGTATCAG 731

QY 661 ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT 720
DB 732 ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT 791
```

```

Db 1872 TAAATTTTAAAGTGTGTATCTTTTTCATCTGGTGTATCTCAGGGTGAATAACGAGCATTC 1931
Qy 1861 ATGGAAGGCGATGCAAGTTTGTTCATTGTGACAGTTTGTGTTTAAATAAACCACATACACACT 1920
Db 1932 ATGGAAGGCGATGCAAGTTTGTTCATTGTGACAGTTTGTGTTTAAATAAACCACATACACACT 1991
Qy 1921 TTATTTAAGATTAAATCTAACTGAAAGTCAAGCTTGGAAATGGACATTTTCCAAAGTATG 1980
Db 1932 TTATTTAAGATTAAATCTAACTGAAAGTCAAGCTTGGAAATGGACATTTTCCAAAGTATG 2051
Qy 1981 TTGTGTAGTACAGATATAAATAAATATGAAATTTCTGATGAGAGTTTTCAGTTTAAATACC 2040
Db 2052 TTGTGTAGTACAGATATAAATAAATATGAAATTTCTGATGAGAGTTTTCAGTTTAAATACC 2111
Qy 2041 AAGTCCTTAGGAGTCTTAACATTTGGCCAGCATCTGTGTTTCAATGACATAAATACGTAA 2100
Db 2112 AAGTCCTTAGGAGTCTTAACATTTGGCCAGCATCTGTGTTTCAATGACATAAATACGTAA 2171
Qy 2101 ACCTATAAGAAATTAAGTTTATTAATTAGGCAATTTATCTGTGTGATAATTTCTTACGGGAG 2160
Db 2172 ACCTATAAGAAATTAAGTTTATTAATTAGGCAATTTATCTGTGTGATAATTTCTTACGGGAG 2231
Qy 2161 AAAGAGATTGATTTGGAAGCAGTTTGGGAAGAAAGTGTCTGTGATAATTTCCAGAAATTT 2220
Db 2232 AAAGAGATTGATTTGGAAGCAGTTTGGGAAGAAAGTGTCTGTGATAATTTCCAGAAATTT 2291
Qy 2221 AATTGATTGGTTACATAAATCTTTTGTGCTTCCAGAA 2256
Db 2292 AATTGATTGGTTACATAAATCTTTTGTGACTTCAGAAA 2327

RESULT 3
US-10-339-656-1
; Sequence 1, Application US/10339656
; Publication No. US20030134319A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-1

Query Match 100.0%; Score 2256; DB 15; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACTCATCCAGTCATTATTATGAAGCAAGGTCTTTGAATGAGCGAGATTATCGGAC 60
Db 72 GACACTCATCCAGTCATTATTATGAAGCAAGGTCTTTGAATGAGCGAGATTATCGGAC 131
Qy 61 CGGAGATCGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTTCTAGACATTAT 120
Db 132 CGGAGATCGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTTCTAGACATTAT 191
Qy 121 CACAGAGATTTGAAGCGGGTATCGAATCCACTCCAGTAAATCTTCAGTCGGAGCAGG 180
Db 192 CACAGAGATTTGAAGCGGGTATCGAATCCACTCCAGTAAATCTTCAGTCGGAGCAGG 251

```

```

Qy 181 AGAAGCAGTCTCTAAAGAGGAGCGCAATAGACACTGTTCAGTCACTAGTCACTGTCGAAG 240
Db 252 AGAAGCAGTCTCTAAAGAGGAGCGCAATAGACACTGTTCAGTCACTAGTCACTGTCGAAG 311
Qy 241 AGCACCAGAGGAAAGATCCAGAGATATAGAGGATGATGAGGAGGTCACCTGATCTGT 300
Db 312 AGCACCAGAGGAAAGATCCAGAGATATAGAGGATGATGAGGAGGTCACCTGATCTGT 371
Qy 301 CAAAGTGGAGAGCTTCTAAAGAGCAAGATATCAAAATCGTGGACACTTTGGGTGAAGAGGCC 360
Db 372 CAAAGTGGAGAGCTTCTAAAGAGCAAGATATCAAAATCGTGGACACTTTGGGTGAAGAGGCC 431
Qy 361 TTTGGCAAAAGTTGTAGAGTGCAATTCATCATGGCATGGATGGCATGTCATGTAGCAGTGAA 420
Db 432 TTTGGCAAAAGTTGTAGAGTGCAATTCATCATGGCATGGATGGCATGTCATGTAGCAGTGAA 491
Qy 421 ATCGTAAAAATGTAGGCGCTTACCGTGAAGCAGCTGTTCAGAAAATCCAAAGTATTAGAG 480
Db 492 ATCGTAAAAATGTAGGCGCTTACCGTGAAGCAGCTGTTCAGAAAATCCAAAGTATTAGAG 551
Qy 481 CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGCCAGATGCTAGAAATGGTTT 540
Db 552 CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGCCAGATGCTAGAAATGGTTT 611
Qy 541 GATCATCATGTCTATGTTTGTGTTTGAATCTACTGGGACTTAGTACTTACGATTTTC 600
Db 612 GATCATCATGTCTATGTTTGTGTTTGAATCTACTGGGACTTAGTACTTACGATTTTC 671
Qy 601 ATTAAGAAAACAGCTTCTGCCAATTTCAAAATGACCAATCAGGAGATGGCTATCAG 660
Db 672 ATTAAGAAAACAGCTTCTGCCAATTTCAAAATGACCAATCAGGAGATGGCTATCAG 731
Qy 661 ATCTGCCAGTCAATAAATTTTATCATATAAATTAACCCATACAGATCTGAAGCCT 720
Db 732 ATCTGCCAGTCAATAAATTTTATCATATAAATTAACCCATACAGATCTGAAGCCT 791
Qy 721 GAAAATATTTTGTGTGAAGTCTGACTATGCTCAAAATATAATTTCTAAAATGAAACGT 780
Db 792 GAAAATATTTTGTGTGAAGTCTGACTATGCTCAAAATATAATTTCTAAAATGAAACGT 851
Qy 781 GATGAACGCACACTGAAAACACAGATATCAAGTTGTGTGACTTTGGAGTGCACAGCTAT 840
Db 852 GATGAACGCACACTGAAAACACAGATATCAAGTTGTGTGACTTTGGAGTGCACAGCTAT 911
Qy 841 GATGATGAACATCAAGTACTTTGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAAT 900
Db 912 GATGATGAACATCAAGTACTTTGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAAT 971
Qy 901 TTGGCTTTAGGTTGGTCTCAGCCTTTGTGATGTTTGGAGCATAGGTTGCATTTTATTGAA 960
Db 972 TTGGCTTTAGGTTGGTCTCAGCCTTTGTGATGTTTGGAGCATAGGTTGCATTTTATTGAA 1031
Qy 961 TATTACCTTTGTTTCAAGTCTTTCAGATCATGATAGTAAAGACACCTGGCAATGATG 1020
Db 1032 TATTACCTTTGTTTCAAGTCTTTCAGATCATGATAGTAAAGACACCTGGCAATGATG 1091
Qy 1021 GAACGAATATTAGGACCCATACCAACACATGATTTCAAGAAAACGAAGAAATGAT 1080
Db 1092 GAACGAATATTAGGACCCATACCAACACATGATTTCAAGAAAACGAAGAAATGAT 1151
Qy 1081 TTTTCAACCAATACCAGTAGATTGGGATGAACACAGTTTCTGCTGTGTAGATATGTTAGGAGA 1140
Db 1152 TTTTCAACCAATACCAGTAGATTGGGATGAACACAGTTTCTGCTGTGTAGATATGTTAGGAGA 1211
Qy 1141 CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCTATGATGAACATGAGAAATCTGTT 1200
Db 1212 CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCTATGATGAAGAAATGAGAAATCTGTT 1271
Qy 1201 GACCTGGTTTCAAGAAATGTTAGAAATATGATCAAACTCAAGAAATTTACCTTGGATGAAGCA 1260
Db 1272 GACCTGGTTTCAAGAAATGTTAGAAATATGATCAAACTCAAGAAATTTACCTTGGATGAAGCA 1331

```

QY 1261 TTGAGCATCTCTTCTTGAATTAAAGAAATGGAATGGAATCAGTGGTCTTAC 1320  
 Db 1332 TTGAGCATCTCTTCTTGAATTAAAGAAATGGAATGGAATCAGTGGTCTTAC 1391  
 QY 1321 TATATACCTCTCTAGAGAGATTACTTAAGACTGTGTCAGTCAATTAACATCTTAATAT 1380  
 Db 1392 TATATACCTCTCTAGAGAGATTACTTAAGACTGTGTCAGTCAATTAACATCTTAATAT 1451  
 QY 1381 TTTTGTAAACATTAATATTTTGTACAGTTAAAGTGAATATTTGTATTTTGTATCAA 1440  
 Db 1452 TTTTGTAAACATTAATATTTTGTACAGTTAAAGTGAATATTTGTATTTTGTATCAA 1511  
 QY 1441 TAGCATTAATTAATTTTGTAAAGCAAGTATGCTTGTATTAATGCAATTAAGAAATTTAAATTT 1500  
 Db 1512 TAGCATTAATTAATTTTGTAAAGCAAGTATGCTTGTATTAATGCAATTAAGAAATTTAAATTT 1571  
 QY 1501 AATTTTCTCTTTTGAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1560  
 Db 1572 AATTTTCTCTTTTGAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1631  
 QY 1561 ATAAATGTGATTTGCTTCTTGTATCATGAGGTCACTCTGAAAGTGAATTTTGTG 1620  
 Db 1632 ATAAATGTGATTTGCTTCTTGTATCATGAGGTCACTCTGAAAGTGAATTTTGTG 1691  
 QY 1621 AGTAAAGGAAATCTTGACTACTTTATATTTCTTAAAGGAATATTTCTTTATATACTTCAA 1680  
 Db 1692 AGTAAAGGAAATCTTGACTACTTTATATTTCTTAAAGGAATATTTCTTTATATACTTCAA 1751  
 QY 1681 TTTAGAACTTAACCTTAAAGTTTCTTCTGTAATTTCTGTAATTTCTGTAATTTCTGTAATTT 1740  
 Db 1752 TTTAGAACTTAACCTTAAAGTTTCTTCTGTAATTTCTGTAATTTCTGTAATTTCTGTAATTT 1811  
 QY 1741 CTCTAGATAAGCAGTACTAGAAACCAAACTTCAGAAATTTTCTGTAATTTCTGTAATTTCT 1800  
 Db 1812 CTCTAGATAAGCAGTACTAGAAACCAAACTTCAGAAATTTTCTGTAATTTCTGTAATTTCT 1871  
 QY 1801 TAAATTTTAAAGTTTGTATTTCTTTTCAATGCGTGAATGTCAGGGTGAATTAACAGACATTC 1860  
 Db 1872 TAAATTTTAAAGTTTGTATTTCTTTTCAATGCGTGAATGTCAGGGTGAATTAACAGACATTC 1931  
 QY 1861 ATGGAAGGCAATGCAATTTGTCATTTGTCAGTGTGACAGTTTGTATTAATAAACCACATACACT 1920  
 Db 1932 ATGGAAGGCAATGCAATTTGTCATTTGTCAGTGTGACAGTTTGTATTAATAAACCACATACACT 1991  
 QY 1921 TTATTTAAGATTTAAATCTAACTGGAAGTCAAGTGTGGAATAATGACATTTTCCAAGTATG 1980  
 Db 1992 TTATTTAAGATTTAAATCTAACTGGAAGTCAAGTGTGGAATAATGACATTTTCCAAGTATG 2051  
 QY 1981 TTTGTTGAGTCAAGATATAAATAAGAAATTTCTGATGAGAGTTTCAAGTTTAAATATACC 2040  
 Db 2052 TTTGTTGAGTCAAGATATAAATAAGAAATTTCTGATGAGAGTTTCAAGTTTAAATATACC 2111  
 QY 2041 AAGTCTTTAGAGTCTTAACATTTGCGCAGCATCTGTTTATCAATGACATAAATACGTAA 2100  
 Db 2112 AAGTCTTTAGAGTCTTAACATTTGCGCAGCATCTGTTTATCAATGACATAAATACGTAA 2171  
 QY 2101 ACCTATAAGATTAAGTTTATTAATTTAGGCAATTTATGCTGTGATTAATTTCTTACGGGAG 2160  
 Db 2172 ACCTATAAGATTAAGTTTATTAATTTAGGCAATTTATGCTGTGATTAATTTCTTACGGGAG 2231  
 QY 2161 AAAGAGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTTCCAGAAATTT 2220  
 Db 2232 AAAGAGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTTCCAGAAATTT 2291  
 QY 2221 AATTGATTTGTTACATAAATCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTT 2256  
 Db 2292 AATTGATTTGTTACATAAATCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTT 2327

RESULT 4

US-10-801-671-1  
 ; Sequence 1, Application US/10801671  
 ; Publication No. US20040152123A1

; GENERAL INFORMATION:  
 ; APPLICANT: YAN, Chunhua et al.  
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
 ; TITLE OF INVENTION: THEREOF  
 ; FILE REFERENCE: CL000758DIV-III  
 ; CURRENT APPLICATION NUMBER: US/10/801,671  
 ; CURRENT FILING DATE: 2004-03-17  
 ; PRIOR APPLICATION NUMBER: 60/227,470  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: 09/810,671  
 ; PRIOR FILING DATE: 2001-03-19  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2354  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-801-671-1

Query Match 100.0%; Score 2256; DB 18; Length 2354;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACTCATCCAGTCAATTTTAAAGCAAGGTCCTTGAATGAGCGAGATTAATCGGAC 60  
 Db 72 GACACTCATCCAGTCAATTTTAAAGCAAGGTCCTTGAATGAGCGAGATTAATCGGAC 131  
 QY 61 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTAT 120  
 Db 132 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTAT 191  
 QY 121 CACAGAGACATTTGAAGCGGTTATCGAATCCACTGCAGTAAATCTTCACTCGGACGAGG 180  
 Db 192 CACAGAGACATTTGAAGCGGTTATCGAATCCACTGCAGTAAATCTTCACTCGGACGAGG 251  
 QY 181 AGAAGCAGTCTTAAAGGAAGCGCAATAGACACTGTTCAAGTCACTAGTCAGTCAGTTCGAAG 240  
 Db 252 AGAAGCAGTCTTAAAGGAAGCGCAATAGACACTGTTCAAGTCACTAGTCAGTCAGTTCGAAG 311  
 QY 241 AGCCACCGAAGAAAGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGT 300  
 Db 312 AGCCACCGAAGAAAGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGT 371  
 QY 301 CAAAGTGGAGAGCTTTAAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGGAGCC 360  
 Db 372 CAAAGTGGAGAGCTTTAAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGGAGCC 431  
 QY 361 TTTGGCAAAAGTTGTAGAGTGCATTTGATCATGGCATGGATGGCATGTCATGTAGCAGTGA 420  
 Db 432 TTTGGCAAAAGTTGTAGAGTGCATTTGATCATGGCATGGATGGCATGTCATGTAGCAGTGA 491  
 QY 421 ATCGTAAATAATGTAGGCGTTACCGTGAACAGCTCGTTCAGAAATCCAGATATTAGAG 480  
 Db 492 ATCGTAAATAATGTAGGCGTTACCGTGAACAGCTCGTTCAGAAATCCAGATATTAGAG 551  
 QY 481 CACTTAAATAGTACTGATCCCAATAGTCTTCCGATGTGTCGATGTGTCGATGTGTCGATGT 540  
 Db 552 CACTTAAATAGTACTGATCCCAATAGTCTTCCGATGTGTCGATGTGTCGATGTGTCGATGT 611  
 QY 541 GATCATCATGTCATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600  
 Db 612 GATCATCATGTCATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 671  
 QY 601 ATTAAGAAACACGCTTTCTGCCATTTTCAATTTGACCAATGACCAATGACCAATGACCAAT 660  
 Db 672 ATTAAGAAACACGCTTTCTGCCATTTTCAATTTGACCAATGACCAATGACCAATGACCAAT 731  
 QY 661 ATCTGCCAGTCAATAAATTTTACATCAATAAATTTTACATCAATAAATTTTACATCAATAA 720  
 Db 732 ATCTGCCAGTCAATAAATTTTACATCAATAAATTTTACATCAATAAATTTTACATCAATAA 791  
 QY 721 GAAAAATTTTGTGTGAGTCTGACTATGCTAGTCAATAATAATTTTCAATAATAAATGAAACGT 780

|    |      |        |  |      |
|----|------|--------|--|------|
| Db | 792  | GA     | AAATATTTTGTGTTGAAAGTCGACTATGTAGTCAAAATATAATCTTAAATGAAACGT  | 851  |
| Qy | 781  | GATGAA | CGCACACTGAAAAAACACAGATATCAAAGTCTGTGACTTTGGAAGTGCAACGTAT    | 840  |
| Db | 852  | GATGAA | CGCACACTGAAAAAACACAGATATCAAAGTCTGTGACTTTGGAAGTGCAACGTAT    | 911  |
| Qy | 841  | GATGAT | GAAACATCACAGTACTTTGGTGTCTACCGGCACCTACAGAGCTCCGAGGTCAAT     | 900  |
| Db | 912  | GATGAT | GAAACATCACAGTACTTTGGTGTCTACCGGCACCTACAGAGCTCCGAGGTCAAT     | 971  |
| Qy | 901  | TTGGCT | TTTAGTTGGTCTCAGCCCTTGATGTGTTGGAGCATAGGTTGCATCTTTATTGAA     | 960  |
| Db | 972  | TTGGCT | TTTAGTTGGTCTCAGCCCTTGATGTGTTGGAGCATAGGTTGCATCTTTATTGAA     | 1031 |
| Qy | 961  | TATTAC | CTTCGGTTTTCACAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGGCAATGATG  | 1020 |
| Db | 1032 | TATTAC | CTTCGGTTTTCACAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGGCAATGATG  | 1091 |
| Qy | 1021 | GAACGA | ATATTAGGACCCATACCAACAACATGATTCAGAAAAACAAGAAAAACGCAAGTAT    | 1080 |
| Db | 1092 | GAACGA | ATATTAGGACCCATACCAACAACATGATTCAGAAAAACAAGAAAAACGCAAGTAT    | 1151 |
| Qy | 1081 | TTTCA  | CAATAACACAGCTAGATTTGGGATGAACAACAGTTCTCGTGGTAGATATGTTAGAGA  | 1140 |
| Db | 1152 | TTTCA  | CAATAACACAGCTAGATTTGGGATGAACAACAGTTCTCGTGGTAGATATGTTAGAGA  | 1211 |
| Qy | 1141 | CGCTG  | CAAAACCGTTTGAAGGAATTTATGCTTTTGTCATCATGAAGAAACATGAGAACTCTTT | 1200 |
| Db | 1212 | CGCTG  | CAAAACCGTTTGAAGGAATTTATGCTTTTGTCATCATGAAGAAACATGAGAACTCTTT | 1271 |
| Qy | 1201 | GACCTG | TTTCGAAGAAATGTTAGAAATATGATCCAACTCAAAGAATTACCTTGGATGAAGCA   | 1260 |
| Db | 1272 | GACCTG | TTTCGAAGAAATGTTAGAAATATGATCCAACTCAAAGAATTACCTTGGATGAAGCA   | 1331 |
| Qy | 1261 | TTGCA  | GCATCCTTTCTTTTGACTTATTAATAAAGAAATGAAATGGGAATCAGTGGTCTTAC   | 1320 |
| Db | 1332 | TTGCA  | GCATCCTTTCTTTTGACTTATTAATAAAGAAATGAAATGGGAATCAGTGGTCTTAC   | 1391 |
| Qy | 1321 | TATAT  | ACTTCTCTAGAGAGATTTACTTTAAGCTGTGTCACTCACTAAAACATTTCTAATAT   | 1380 |
| Db | 1392 | TATAT  | ACTTCTCTAGAGAGATTTACTTTAAGCTGTGTCACTCACTAAAACATTTCTAATAT   | 1451 |
| Qy | 1381 | TTTTG  | TAAAACATTTAAATTTATTTGTACAGTTTAAAGTGAATATTTGTATGTTTGTATCAA  | 1440 |
| Db | 1452 | TTTTG  | TAAAACATTTAAATTTATTTGTACAGTTTAAAGTGAATATTTGTATGTTTGTATCAA  | 1511 |
| Qy | 1441 | TAGCAT | AATTAACTTTTAAGCAAGTATGTTGTATAATGCAATTAGAAAAATTTAAATTT      | 1500 |
| Db | 1512 | TAGCAT | AATTAACTTTTAAGCAAGTATGTTGTATAATGCAATTAGAAAAATTTAAATTT      | 1571 |
| Qy | 1501 | AAATTT | CTTTTTCGAAATACCATTTTTAAATACCTTTTGAATAATCCTTTGTGTCCAGTG     | 1560 |
| Db | 1572 | AAATTT | CTTTTTCGAAATACCATTTTTAAATACCTTTTGAATAATCCTTTGTGTCCAGTG     | 1631 |
| Qy | 1561 | ATAAT  | CTGTATGATCTTTGCTTTGTACATGGAGGTCACTCTCGAAGTGAATTTTTTTTG     | 1620 |
| Db | 1632 | ATAAT  | CTGTATGATCTTTGCTTTGTACATGGAGGTCACTCTCGAAGTGAATTTTTTTTG     | 1691 |
| Qy | 1621 | AGTAAA | AGGAAATCTTGACTACTTTATATCTTTAAAGGAATATTTCTTTATATACTTCAAA    | 1680 |
| Db | 1692 | AGTAAA | AGGAAATCTTGACTACTTTATATCTTTAAAGGAATATTTCTTTATATACTTCAAA    | 1751 |
| Qy | 1681 | TTTAGA | ACTTAACTTTTAAAGTTTTCTTCTGTATTAATTTGTTGAACGGGTGATTATTTAA    | 1740 |
| Db | 1752 | TTTAGA | ACTTAACTTTTAAAGTTTTCTTCTGTATTAATTTGTTGAACGGGTGATTATTTAA    | 1811 |
| Qy | 1741 | CTCT   | TAGATAGCAGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTAT    | 1800 |
| Db | 1812 | CTCT   | TAGATAGCAGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTAT    | 1871 |
| Qy | 1801 | TAAAT  | TTTTTAAGTGTGTATCTTTTTTCAATTGGGTGATGTCAGGGTGATAACCAACATTC   | 1860 |

|      |   |      |
|------|---|------|
| 1872 | TAAATTTTAAGTGTGTAATCTCTTTTTCATTTGGGTGATGTCAGGGTGATAACCGACATTC   | 1931 |
| 1861 | ATGGAAGGCATGCAGTTGTGTCATTTGTGACAGTTGTGTTTAATAAAAAACCATACACACT   | 1920 |
| 1932 | ATGGAAGGCATGCAGTTGTGTCATTTGTGACAGTTGTGTTTAATAAAAAACCATACACACT   | 1991 |
| 1921 | TTATTTTAAGATTAAATCTAACTCGGAAGTCAGCTTGGAAAATGGACATTTTCCAGTATG    | 1980 |
| 1992 | TTATTTTAAGATTAAATCTAACTCGGAAGTCAGCTTGGAAAATGGACATTTTCCAGTATG    | 2051 |
| 1981 | TTTGGTCAGTCACAGATATATAAAATAGAAATCTTGATGAGAGGTTTCAGTTTAAAAATACC  | 2040 |
| 2052 | TTTGGTCAGTCACAGATATATAAAATAGAAATCTTGATGAGAGGTTTCAGTTTAAAAATACC  | 2111 |
| 2041 | AAGTCCTTAGGAGCTTTAAACATTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAA  | 2100 |
| 2112 | AAGTCCTTAGGAGCTTTAAACATTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAA  | 2171 |
| 2101 | ACCTATAAGAAATTAAGTTTATTAAATAGGCAAAATTAATGCTGCTGATTAATCTTACGGGAG | 2160 |
| 2172 | ACCTATAAGAAATTAAGTTTATTAAATAGGCAAAATTAATGCTGCTGATTAATCTTACGGGAG | 2231 |
| 2161 | AAAGAGCATTTGATTTGGAAAACAGTTTGGGAAGAAAGTCTGCTGAAATTTTCCGAAATTT   | 2220 |
| 2232 | AAAGAGCATTTGATTTGGNAAGCAGTTGGGAAGAAAGTCTGCTGAAATTTTCCGAAATTT    | 2291 |
| 2221 | AATTGATTTGGTTACATAAACCTTTTGTAGCTTCAGAAA                         | 2256 |
| 2292 | AATTGATTTGGTTACATAAACCTTTTGTAGCTTCAGAAA                         | 2327 |

RESULT 5  
US-10-425-114-26852  
; Sequence 26852, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 26852  
; LENGTH: 4035  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB4654-056-G4\_FLI  
US-10-425-114-26852

|    | Query Match                 | 84.6%   | Score 1908.4 | DB 17    | Length 4035 |
|----|-----------------------------|---|--------------|----------|-------------|
|    | Best Local Similarity 99.9% | Pred. No. 0   |              |          |             |
|    | Matches 1909                | Conservative 0  | Mismatches 1 | Indels 0 | Gaps 0      |
| QY | 327                         | ATATGAANAATCGTGGACACTTTTGGGTGAAGAGAGCCTTTGGCAAAAGTTGTAGAGTGCATTGA | 386          |          |             |
| Db | 2126                        | AGATGAANAATCGTGGACACTTTTGGGTGAAGAGAGCCTTTGGCAAAAGTTGTAGAGTGCATTGA | 2185         |          |             |
| QY | 387                         | TCATGGCATGGATGGCATGTCATGTAGCAGTGTAAAAATCGTAAAAAAATGTAGCGCGTTACCG  | 446          |          |             |
| Db | 2186                        | TCATGGCATGGATGGCATGTCATGTAGCAGTGTAAAAATCGTAAAAAAATGTAGCGCGTTACCG  | 2245         |          |             |
| QY | 447                         | TGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTACTGATCCCAATAG   | 506          |          |             |
| Db | 2246                        | TGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTACTGATCCCAATAG   | 2305         |          |             |
| QY | 507                         | TGTCCTCCGAGTGTCCAGATGCTAGAAATGGTTTGTATCATCATGGTCATCGTTGTGTATTGT   | 566          |          |             |

Db 2306 TGTCTTCCGATGTGTCAGATGCTAGATGGTTTGATCATCATGTGTCAATGTTGTTATGTT 2365  
QY 567 GTTTGAACTACTCGGACCTTAGTACTTACGATTTCAATAAGAAACACAGCTTCTGCGCAAT 626  
Db 2366 GTTTGAACTACTCGGACCTTAGTACTTACGATTTCAATAAGAAACACAGCTTCTGCGCAAT 2425  
QY 627 TCAATTCACCACTACGAGCAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTCACA 686  
Db 2426 TCAATTCACCACTACGAGCAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTCACA 2485  
QY 687 TCATAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGTGTGAGTCTGA 746  
Db 2486 TCATAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGTGTGAGTCTGA 2545  
QY 747 CTATGTAGTCAAAATATAATTTCTAATAATGAACCTGATGAACGCGACACTGAAACACACAGA 806  
Db 2546 CTATGTAGTCAAAATATAATTTCTAATAATGAACCTGATGAACGCGACACTGAAACACACAGA 2605  
QY 807 TATCAAAAGTTGTGACTTTGGAAGTGCAACGATGATGATGAACATCAAGTACTTTGGT 866  
Db 2606 TATCAAAAGTTGTGACTTTGGAAGTGCAACGATGATGATGAACATCAAGTACTTTGGT 2665  
QY 867 GTCTACCCGCACTACAGACTCCGAGGTCAATTTGGCTTTAGTGGTCTCAGCCTTG 926  
Db 2666 GTCTACCCGCACTACAGACTCCGAGGTCAATTTGGCTTTAGTGGTCTCAGCCTTG 2725  
QY 927 TGATGTTTGGAGCATAGGTTGCAATCTTATTAATTAATTAATTAATTAATTAATTAATTAAT 986  
Db 2726 TGATGTTTGGAGCATAGGTTGCAATCTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2785  
QY 987 GACTCATGATAGTAAGAGCACCTGGCAATGATGAACGAATATTTAGGACCCATACCA 1046  
Db 2786 GACTCATGATAGTAAGAGCACCTGGCAATGATGAACGAATATTTAGGACCCATACCA 2845  
QY 1047 ACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACATTAACCCATACCA 1106  
Db 2846 ACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACATTAACCCATACCA 2905  
QY 1107 TGAACACAGTTCTGCTGTTGATATGTTAGAGAGCGCTGCAACCGTTGAAGGAATTTAT 1166  
Db 2906 TGAACACAGTTCTGCTGTTGATATGTTAGAGAGCGCTGCAACCGTTGAAGGAATTTAT 2965  
QY 1167 GCTTTGTCATGATGAAGAACATAGAAAACTGTTTGACCTGTTGCGAAGATGTTAGAATA 1226  
Db 2966 GCTTTGTCATGATGAAGAACATAGAAAACTGTTTGACCTGTTGCGAAGATGTTAGAATA 3025  
QY 1227 TGATCCAACTCAAGAAATTAACCTGGATGAAGCATTCAGCATCTCTTTCTTGACTTAT 1286  
Db 3026 TGATCCAACTCAAGAAATTAACCTGGATGAAGCATTCAGCATCTCTTTCTTGACTTAT 3085  
QY 1287 AAAAAAGAAATGAATGGAAATCAGTGTCTTACTATATCTCTCTAGAGAGATTA 1346  
Db 3086 AAAAAAGAAATGAATGGAAATCAGTGTCTTACTATATCTCTCTAGAGAGATTA 3145  
QY 1347 TAAGACTGTGTCAGTCAACTAAACATCTTAATTTTGTAAACATTAATTTTGTGA 1406  
Db 3146 TAAGACTGTGTCAGTCAACTAAACATCTTAATTTTGTAAACATTAATTTTGTGA 3205  
QY 1407 CAGTTAAGTGAATATTTGATGTTTGTGATCAATPAGCATTAATTAATTTGTAAGCAAGT 1466  
Db 3206 CAGTTAAGTGAATATTTGATGTTTGTGATCAATPAGCATTAATTTGTAAGCAAGT 3265  
QY 1467 ATGGTCTTGATAATGCATTAAGAAAAATTAATTAATTTTCTTTTTCGAAATTAACATTT 1526  
Db 3266 ATGGTCTTGATAATGCATTAAGAAAAATTAATTAATTTTCTTTTTCGAAATTAACATTT 3325  
QY 1527 TTAATACCTTTGAAATATCTTTGTCAGTGAATAATGATGATCTTTGCTCTTTG 1586  
Db 3326 TTAATACCTTTGAAATATCTTTGTCAGTGAATAATGATGATCTTTGCTCTTTG 3385  
QY 1587 TACATGGAGGTCACTCTGAAAGTATTTTGTGATGAATAAGGAAATCTTGACTACTTTTA 1646

Db 3386 TACATGGAGGTCACTCTGAAGTGAATTTTGTGATGAAGAAATCTTGACTACTTTA 3445  
QY 1647 TATCTTAAAGGAATATTTCTTTATATATCTCAAAATTTAGAACTTAACCTTTAAAGCTTTT 1706  
Db 3446 TATCTTAAAGGAATATTTCTTTATATATCTCAAAATTTAGAACTTAACCTTTAAAGCTTTT 3505  
QY 1707 CTTCTGTAATTTGTTGAACGGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1766  
Db 3506 CTTCTGTAATTTGTTGAACGGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3565  
QY 1767 AAAAATCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTAATTAATTAATTAATTAATTAAT 1826  
Db 3566 AAAAATCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTAATTAATTAATTAATTAATTAAT 3625  
QY 1827 CATTTGGGTGATGTCAGGGTGATAACACAGACATTCATGAAAGGCCATGACAGTTTGTCCATTT 1886  
Db 3626 CATTTGGGTGATGTCAGGGTGATAACACAGACATTCATGAAAGGCCATGACAGTTTGTCCATTT 3685  
QY 1887 GTGACAGTTTGTTAATAAACCACATACACACTTTTATTAAGATTAAATCTTAACCTGGA 1946  
Db 3686 GTGACAGTTTGTTAATAAACCACATACACACTTTTATTAAGATTAAATCTTAACCTGGA 3745  
QY 1947 AAGTCAGCTTGGAAAAATGGACATTTCCAAAGTATGTTTGGTGAGTCACAGATATAAAAAATA 2006  
Db 3746 AAGTCAGCTTGGAAAAATGGACATTTCCAAAGTATGTTTGGTGAGTCACAGATATAAAAAATA 3805  
QY 2007 GAAATTTCTGATGAGAGGTTTTCAGTTTTTAAATACCAAGTCTTTAGGAGTCTTAAACATTTGGC 2066  
Db 3806 GAAATTTCTGATGAGAGGTTTTCAGTTTTTAAATACCAAGTCTTTAGGAGTCTTAAACATTTGGC 3865  
QY 2067 CAGCATCTGTTTATCAAAATGACATAAATACGTAAACCTATTAAGAAATTAAGTTTATTAATTT 2126  
Db 3866 CAGCATCTGTTTATCAAAATGACATAAATACGTAAACCTATTAAGAAATTAAGTTTATTAATTT 3925  
QY 2127 AGGCAATTTATGCTGTGATTAATTTCTTACCGGAGAAAGAGGATTTGATTGGAAGCAGTT 2186  
Db 3926 AGGCAATTTATGCTGTGATTAATTTCTTACCGGAGAAAGAGGATTTGATTGGAAGCAGTT 3985  
QY 2187 TGGGAAGAAAGTGTCTGTGATAATTTCCAGAAATTTAAATTTGATTGTTTACAT 2236  
Db 3986 TGGGAAGAAAGTGTCTGTGATAATTTCCAGAAATTTAAATTTGATTGTTTACAT 4035

## RESULT 6

US-10-267-502-135  
; Sequence 135: Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; APPLICANT: Galant, Ron  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 135  
; LENGTH: 1446  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-267-502-135

Query Match 57.0%; Score 1286; DB 17; Length 1446;

Best Local Similarity 100.0%; Pred. No. 2.8e-260;

Matches 1286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GTCATTATTAGAACGAAGTCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 73

Db 161 GTCATTATTAGAACGAAGTCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 220

QY 74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAATCACAGAGCAT 133

Db 221 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAATCACAGAGCAT 280

|    |      |   |      |
|----|------|---|------|
| QY | 134  | AAAGCGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGACGAGGAGACGAGTCTTA     | 193  |
| DB | 281  | AAAGCGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGACGAGGAGACGAGTCTTA     | 340  |
| QY | 194  | AAAGGAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTCCGAAGGCCACCGAAGGA     | 253  |
| DB | 341  | AAAGGAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTCCGAAGGCCACCGAAGGA     | 400  |
| QY | 254  | AAAGATCCAGGAGTATAGAGATGATGAGGAGGGTCACTTGATCTGTCAAAGTCGAGACG     | 313  |
| DB | 401  | AAAGATCCAGGAGTATAGAGATGATGAGGAGGGTCACTTGATCTGTCAAAGTCGAGACG     | 460  |
| QY | 314  | TTCTAAGAGCAAGATATGAATCTGTGACACTTTTGGGTGAAGAGCCCTTTGGCAAAAGTTG   | 373  |
| DB | 461  | TTCTAAGAGCAAGATATGAATCTGTGACACTTTTGGGTGAAGAGCCCTTTGGCAAAAGTTG   | 520  |
| QY | 374  | TAGAGTCGATTTGATCATGGCGATGATGGCATGCGATGTAGCAGTCGAAATCGTAAAAAATG  | 433  |
| DB | 521  | TAGAGTCGATTTGATCATGGCGATGATGGCATGCGATGTAGCAGTCGAAATCGTAAAAAATG  | 580  |
| QY | 434  | TAGCCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAGTATTAGAGCACCTTAAATAGTA  | 493  |
| DB | 581  | TAGCCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAGTATTAGAGCACCTTAAATAGTA  | 640  |
| QY | 494  | CTGATCCCAATAGTGTCTTCGATGTGTCAGAGTGCTAGAAATGGTTTGATCATCATGGTC    | 553  |
| DB | 641  | CTGATCCCAATAGTGTCTTCGATGTGTCAGAGTGCTAGAAATGGTTTGATCATCATGGTC    | 700  |
| QY | 554  | ATGTTTGTATTGTGTTTGAACCTACTGGGACTTTAGTACTTACGATTTCAATTAAGAAAAACA | 613  |
| DB | 701  | ATGTTTGTATTGTGTTTGAACCTACTGGGACTTTAGTACTTACGATTTCAATTAAGAAAAACA | 760  |
| QY | 614  | GCTTTCTGCCAATTCAAATTGACACATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA    | 673  |
| DB | 761  | GCTTTCTGCCAATTCAAATTGACACATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA    | 820  |
| QY | 674  | TAAATTTTTTACATCATATAAATTAACCCATACAGATCTCGAAGCTCGAAATATTTTGT     | 733  |
| DB | 821  | TAAATTTTTTACATCATATAAATTAACCCATACAGATCTCGAAGCTCGAAATATTTTGT     | 880  |
| QY | 734  | TTGTGAAGTCTGACTATGTAGTCAAAATATAATTCCTAAAAAGAAACGCGATGAAACGAC    | 793  |
| DB | 881  | TTGTGAAGTCTGACTATGTAGTCAAAATATAATTCCTAAAAAGAAACGCGATGAAACGAC    | 940  |
| QY | 794  | TGAAAAACACAGATATCAAGTCTGCTACCTTTGGAAGTGCACCTATGATGATCAATCA      | 853  |
| DB | 941  | TGAAAAACACAGATATCAAGTCTGCTACCTTTGGAAGTGCACCTATGATGATCAATCA      | 1000 |
| QY | 854  | ACAGTACTTTGGTGTCTACCCGGCACATACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT   | 913  |
| DB | 1001 | ACAGTACTTTGGTGTCTACCCGGCACATACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT   | 1060 |
| QY | 914  | GGTCTCAGCCCTTGATGTTTGGAGCATAGGTTGCAATTCCTTAATGAAATATTACCTTGGTT  | 973  |
| DB | 1061 | GGTCTCAGCCCTTGATGTTTGGAGCATAGGTTGCAATTCCTTAATGAAATATTACCTTGGTT  | 1120 |
| QY | 974  | TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATCGATGGAACGAAATATTAG | 1033 |
| DB | 1121 | TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATCGATGGAACGAAATATTAG | 1180 |
| QY | 1034 | GACCCATACCAACAACATGATTTCAGAAAAACGAAAAACGCAAGTATTTTCCACCAATACC   | 1093 |
| DB | 1181 | GACCCATACCAACAACAATGATTTCAGAAAAACGAAAAACGCAAGTATTTTCCACCAATACC  | 1240 |
| QY | 1094 | AGCTTAGATTGGGATGAACACAGTTCTGCTGCTGATATGTTTAGAGACGCTGCAAAACCGT   | 1153 |
| DB | 1241 | AGCTTAGATTGGGATGAACACAGTTCTGCTGCTGATATGTTTAGAGACGCTGCAAAACCGT   | 1300 |
| QY | 1154 | TGAAGGAATTTATGCTTTGTTCATGATGAAGAAACATGAAAACTGTTTGACCTGGTTCCGAA  | 1213 |
| DB | 1301 | TGAAGGAATTTATGCTTTGTTCATGATGAAGAAACATGAAAACTGTTTGACCTGGTTCCGAA  | 1360 |

|   |      |   |      |
|---|------|---|------|
| Qy  | 1214 | GAATGTTAGAAATATGATCCAACTCAAGAATTAACCTTGGATGAAGCAATTGCAGACATCCTT | 1273 |
| Db  | 1361 | GAATGTTAGAAATATGATCCAACTCAAGAATTAACCTTGGATGAAGCAATTGCAGACATCCTT |      |
| Qy  | 1274 | TCCTTGACTCTATTAAAAAGAAATGA                                      | 1299 |
| Db  | 1421 | TCCTTGACTCTATTAAAAAGAAATGA                                      | 1446 |
| RESULT 7  |      |   |      |
| US-10-641-643-699   |      |   |      |
| ; Sequence 699, Application US/10641643                           |      |   |      |
| ; Publication No. US20040077003A1                                 |      |   |      |
| ; GENERAL INFORMATION:  |      |   |      |
| ; APPLICANT: Cocks, Benjamin G.                                   |      |   |      |
| ; Susan G. Stuart   |      |   |      |
| ; Jeffrey J. Seilhamer  |      |   |      |
| ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL |      |   |      |
| ; GENE EXPRESSION   |      |   |      |
| ; NUMBER OF SEQUENCES: 1508                                       |      |   |      |
| ; CORRESPONDENCE ADDRESS:   |      |   |      |
| ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.                         |      |   |      |
| ; STREET: 3174 PORTER DRIVE                                       |      |   |      |
| ; CITY: PALO ALTO   |      |   |      |
| ; STATE: CALIFORNIA   |      |   |      |
| ; COUNTRY: USA  |      |   |      |
| ; ZIP: 94304  |      |   |      |
| ; COMPUTER READABLE FORM:   |      |   |      |
| ; MEDIUM TYPE: Floppy disk  |      |   |      |
| ; COMPUTER: IBM PC compatible                                     |      |   |      |
| ; OPERATING SYSTEM: PC-DOS/MS-DOS                                 |      |   |      |
| ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2               |      |   |      |
| ; CURRENT APPLICATION DATA:                                       |      |   |      |
| ; APPLICATION NUMBER: US/10/641,643                               |      |   |      |
| ; FILING DATE: 14-Aug-2003  |      |   |      |
| ; CLASSIFICATION: <Unknown>                                       |      |   |      |
| ; PRIOR APPLICATION DATA:   |      |   |      |
| ; APPLICATION NUMBER: <Unknown>                                   |      |   |      |
| ; FILING DATE: <Unknown>  |      |   |      |
| ; ATTORNEY/AGENT INFORMATION:                                     |      |   |      |
| ; NAME: Zeller, Karen J.  |      |   |      |
| ; REGISTRATION NUMBER: 37,071                                     |      |   |      |
| ; REFERENCE/DOCKET NUMBER: PA-0001 US                             |      |   |      |
| ; TELECOMMUNICATION INFORMATION:                                  |      |   |      |
| ; TELEPHONE: (650) 855-0555                                       |      |   |      |
| ; TELEFAX: (650) 845-4166   |      |   |      |
| ; INFORMATION FOR SEQ ID NO: 699:                                 |      |   |      |
| ; SEQUENCE CHARACTERISTICS:                                       |      |   |      |
| ; LENGTH: 1456 base pairs   |      |   |      |
| ; TYPE: nucleic acid  |      |   |      |
| ; STRANDEDNESS: single  |      |   |      |
| ; TOPOLOGY: linear  |      |   |      |
| ; IMMEDIATE SOURCE:   |      |   |      |
| ; LIBRARY: HNT2AGT01  |      |   |      |
| ; CLONE: 488842   |      |   |      |
| ; SEQUENCE DESCRIPTION: SEQ ID NO: 699 :                          |      |   |      |
| US-10-641-643-699   |      |   |      |
| Query Match 52.2%; Score 1178.2; DB 17; Length 1456;              |      |   |      |
| Best Local Similarity 99.7%; Pred. No. 1.4e-237;                  |      |   |      |
| Matches 1191; Conservative 0; Mismatches 3; Indels 1; Gaps 1;     |      |   |      |
| Qy  | 932  | TTTGGAGCATAGGTTGCATCTTATTGAATATTACCTTGGTTTTACAGTCTTTTCAGACTC    | 991  |
| Db  | 262  | TGTAGCCCATAGGTTGCATCTTATTGAATATTACCTTGGTTTTACAGTCTTTTCAGACTC    | 321  |
| Qy  | 992  | ATGATAGTAAGAGCACTCGCAATGATGAACGAATATTAGGACCCCATACCAACACA        | 1051 |
| Db  | 322  | ATGATAGTAAGAGCACTCGCAATGATGAACGAATATTAGGACCCCATACCAACACA        | 381  |
| Qy  | 1052 | TGATTCAGAAAAACAAGAAAAACGAAGTATTTTTCACATAACAGCTAGATTGGATGAAC     | 1111 |
| Db  | 382  | TGATTCAGAAAAACAAGAAAAACGAAGTATTTTTCACATAACAGCTAGATTGGATGAAC     | 441  |





Db 899 TAAATTTTACATCATATAAATTAACACACACCGGACCTAAACCTGAAAAATATTTTAT 958  
Qy 734 TTGTGAGTCTGACTATGCTCAATATAATCTTAAATGAACGCTGAACGACAC 793  
Db 959 TTGTGAAGTCTGACTATGCTCAATATAATCTTAAATGAACGCTGAACGACAC 1018  
Qy 794 TGAATAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACCGTATGATGATCAACATC 853  
Db 1019 TGAATAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACCATATGACGACGACATC 1078  
Qy 854 ACAGTACTTTGGTGTCTACCCGGCCTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 913  
Db 1079 ATAGTACTTTGGTGTCTACCAAGGCACTACAGGCTCCAGAGGTCAATTTTGGCTCTAGGTT 1138  
Qy 914 GGTCTCAGCCTTTGATGCTTTGGAGCATAGTTGCAATCTTATTCAATATTACCTTGGTT 973  
Db 1139 GGTCTCAGCCTTTGATGCTTTGGAGCATAGGCTGCAATCTTATTGAGTACTACCTTGGGT 1198  
Qy 974 TCAGAGTCTTTCAAGTCTCATATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAG 1033  
Db 1199 TCAGAGTCTTTCAAGTCTCATATAGTAAAGAGCACCTGGCAATGATGGAACGATCTTAG 1258  
Qy 1034 GACCCATACCAACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACATAACC 1093  
Db 1259 GACCCATACCAACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACATAACC 1318  
Qy 1094 AGCTAGATGGGATGAACACAGTCTCTGCTGGTGTAGATATGTTAGGAGACGCTGCAACCGT 1153  
Db 1319 AGCTAGATGGGATGAACACAGTCTCTGCTGGGAGATATGTTAGGAGACGCTGCAACCGT 1378  
Qy 1154 TGAAGGAATTTATGCTTTGCTCATGATGAAGAACATGAGAAAACTGTTTGACCTGGTTGAA 1213  
Db 1379 TAAAGGAATTTATGCTGCTGCTCATGACGAGAGCATGAGAGCTGTTTGACCTGGTTGAA 1438  
Qy 1214 GAATGTTAGATATGATCCAACTCAAGAAATACCTTGGATGAAGCATTTGAGCATTCCTT 1273  
Db 1439 GAATGTTGAGTATGACCCAGCAGAGGATCACCTTGGATGAAGCATTTGAGCATTCCTT 1498  
Qy 1274 TCTTTGACTTATTAAGAAAGAAATGAATGG 1304  
Db 1499 TCTTTGACTTATTAAGAAAGAAATGAGTGGG 1529

RESULT 9  
US-10-825-177-26  
; Sequence 26, Application US/10825177  
; Publication No. US20040259220A1  
; GENERAL INFORMATION:  
; APPLICANT: NAYLER, Oliver  
; APPLICANT: ULLRICH, Axel  
; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS  
; FILE REFERENCE: 038602/0431  
; CURRENT APPLICATION NUMBER: US/10/825,177  
; CURRENT FILING DATE: 2004-04-16  
; PRIOR APPLICATION NUMBER: US/09/905,999  
; PRIOR FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: 09/127,248  
; PRIOR FILING DATE: 1999-07-31  
; PRIOR APPLICATION NUMBER: PCT/IB97/00946  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 08/877,150  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 60/034,286  
; PRIOR FILING DATE: 1996-12-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 26  
; LENGTH: 1549  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-825-177-26

Query Match 48.6%; Score 1097.4; DB 18; Length 1549;  
Best Local Similarity 90.6%; Pred. No. 1.4e-220;  
Matches 1170; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
Qy 14 GTCAATTTTAGAACGAGTCCCTTGAATGAGCGAGATATTCGGGACCGGAGATACGTTG 73  
Db 239 GTCACTATTTAGAACGAGATGCTTGAATGAGAGATATTCGGGACCGGAGATACATTTG 298  
Qy 74 ACGAATACAGGAATGACTTCTGTGAAGATATGTTCTCTAGACATTTATCACAGAGACATTTG 133  
Db 299 ATGAATACAGGAATGACTTCTGTGAAGATATGTTCTCAAGACATTTACCATAGAGCGTTG 358  
Qy 134 AAACGGGTATPCGAATCCATGTCAGTAAATCTTCAGTCCGAGCAGAGAGAGTCTCTTA 193  
Db 359 AAAGCACTTACCGGATCCATTCAGTAAATCTTCAGTCCGAGCAGAGAGAGTCTCTTA 418  
Qy 194 AAAGGAGCGCAATAGACATCTGTTCAAGTCAATCAGTCCGTTCCGAAGCCACCGAAGGA 253  
Db 419 AGAAGGAGCGTAAATAGACCTTGTCAAGTCAATCAGTCCGATTCGAAGAGCCACCGAAGGA 478  
Qy 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTCTGATCTGTCAAAAGTGAGACG 313  
Db 479 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTCTGATCTGTCAAAAGTGAGACG 538  
Qy 314 TTCTAAGAGCAAGATATGAATCTGTGGACACTTTGGGTGAAGAGCCCTTTGGCAAGTTG 373  
Db 539 TTCTAAGAGCAAGATATGAATCTGTGGACACTTTTAGGTGAAGAGCCCTTTGGCAAGTTG 598  
Qy 374 TAGAGTGCATTTGATGCGGATGATGCGATGCGATGAGCAGTGAAGTCAAAATCTTAAAGAAATG 433  
Db 599 TAGAGTGCATTTGATGCGGATGATGCGCTTACATGATGAGCAGTGAAGTCAAAATCTTAAAGAAATG 658  
Qy 434 TAGGCGGTTTACCGTGAAGAGCAGTCTGTTTCAAGAAATCCAAGTATTAGAGCACTTAAATAGTA 493  
Db 659 TAGGCGGTTTACCGGAGGAGCAGTCTGTTTCAAGAAATCCAAGTATTAGAGCACTTGAACGCA 718  
Qy 494 CTGATCCCAATAGTGTCTTCGATGCTGTCAGATGCTGATGATGTTGATCATCATGTC 553  
Db 719 CTGACCCCAACAGTGTCTTCGATGCTGTCAGATGCTGATGATGTTGATCATCATGTC 778  
Qy 554 ATGTTTGTATTTGCTTTGAACTTCTGGGACTTAGTACTTACGATTTTCAATTAAGAAAAACA 613  
Db 779 ATGTTTGTATTTGCTTTGAGCTGCTGGGACTTAGTACTTACGATTTTCAATTAAGAAAAACA 838  
Qy 614 GCTTTCTGCCATTTCAAAATTTGACCAATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 673  
Db 839 GTTTTCTGCCATTTCAAAATTTGATCAGATCAGGCAAAATGGCTTATCAGATCTGCCAGTCTA 898  
Qy 674 TAAATTTTTHACATCATATAAATTAACCCATACAGATCTCGAAGCTGAAAAATATTTTGT 733  
Db 899 TAAATTTTTHACATCATATAAATTAACCCATACAGATCTCGAAGCTGAAAAATATTTTAT 958  
Qy 734 TTGTGAGTCTGATGCTGATGCTCAAAATATAATTTCTTAAATGAAAGCTGATGAACACAC 793  
Db 959 TTGTGAGTCTGATGCTGATGCTCAAAATATAATTTCTTAAATGAAAGCTGATGAACACAC 1018  
Qy 794 TGAATAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACCGTATGATGATGAACATC 853  
Db 1019 TGAATAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACCATATGACGACGACATC 1078  
Qy 854 ACAGTACTTTGGTGTCTACCCGGCCTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 913  
Db 1079 ATAGTACTTTGGTGTCTACCAAGGCACTACAGGCTTCCAGAGGTCAATTTTGGCTCTAGGTT 1138  
Qy 914 GGTCTCAGCCTTTGATGCTTTGGAGCATAGTGTGCAATCTTATTGAATATTACCTTGGTT 973  
Db 1139 GGTCTCAGCCTTTGATGCTTTGGAGCATAGGCTGCAATCTTATTGAGTACTACCTTGGTT 1198  
Qy 974 TCAGAGTCTTTCAAGTCTCATATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAG 1033  
Db 1199 TCAGAGTCTTTCAAGTCTCATATAGTAAAGAGCACCTGGCAATGATGGAACGATCTTAG 1258  
Qy 1034 GACCCATACCAACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACATAACC 1093



```
Db 1259 GACCCATCCGACACATATGATCCAGAGACCAAGGAACCGCAAGTATTTCCACCAATAACC 1318
QY 1094 AGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1153
Db 1319 AGCTAGATTGGGACGAGCATAGTTCAGCTGGGAGATATGTTAGGAGACGCTGCAAGCCGT 1378
QY 1154 TGAAGAAATTTATGCTTTGCTATGATGAAGAAACATGAGAAACCTGTTGACCTGGTTCGAA 1213
Db 1379 TAAAGAAATTTATGCTGCTGATGACCAAGAGCATGAGAAGCTGTTGACCTGGTTCGAA 1438
QY 1214 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGAGGATCCCTT 1273
Db 1439 GAATGTTGAGATGATGACCCAGGAGAGGATCACCTTGGATGAAGCATTTGAGGATCCCTT 1498
QY 1274 TCTTTGACTTATTAAGAAAGAAATGAATGG 1304
Db 1499 TCTTTGACTTATTAAGAAAGAAATGATGGG 1529

RESULT 10
US-10-267-502-138
; Sequence 138, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 138
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-267-502-138

Query Match 48.6%; Score 1097.2; DB 17; Length 1446;
Best Local Similarity 90.8%; Pred. No. 1.5e-220;
Matches 1168; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 14 GTCATTTATTAGAACCAAGCTCTTCAATGACGAGATTTATCGGACCGGAGATACGTTG 73
Db 161 GTCATTTATTAGAACCAAGATGCTTGAATGAGAGATTTATCGGACCGGAGATACATTG 220
QY 74 AGCAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCAGAGACATTG 133
Db 221 ATGAATACAGAAATGACTACTCGGAAGGATATGTTCTCAAGACATTTACCATAGAGACGTTG 280
QY 134 AAAGCGGTATCGAATCCACTCCAGTAAATCTTCAGTCCGACGAGGAGAGACGTCCTTA 193
Db 281 AAAGCACTTACCGGATCCATTCAGTAAATCTTCAGTCCGACGAGGAGAGACGCCCTTA 340
QY 194 AAAGGAAGCGCAATAGACACTGTTCAAGTCACTCAGTCAGTTCGAGAGACCCACCGAAGGA 253
Db 341 AGAGAAAGCGTAATAGACCCCTGTCAAGTCACTCAGTCGCTTCGAAGAGACCCACCGAAGGA 400
QY 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCMAAGTGAGAGCG 313
Db 401 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCMAAGTGAGAGCG 460
QY 314 TTCTAAGAGCAGATATGAAATCGTGGACACATTTGGGTGAAGGAGCCTTTGCAAGTTG 373
Db 461 TTCTAAGAGCAGATATGAAATCGTGGACACATTTAGGTGAAGGAGCCTTTGCAAGTTG 520
QY 374 TAGAGTGCAATTGATCATGCGCATGGATGCGATGATGATGATGATGATGATGATGATGATG 433
Db 521 TAGAGTGCAATTGATCATGCGCATGGATGCGCTTACATGATGATGATGATGATGATGATG 580
QY 434 TAGGCGGTGTACCGTGAAGCAGCTGTTTCAGAAATCCAAAGTATTTAGAGCACTTAAATAGTA 493
```

## RESULT 11

US-09-810-671-3

; Sequence 3, Application US/09810671

; Publication No. US20020076783A1

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; FILE REFERENCE: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE OF INVENTION: THEREOF

; FILE REFERENCE: CL000758

; CURRENT APPLICATION NUMBER: US/09/810,671

; CURRENT FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

```
Db 581 TAGGACGTTACCGGAGGAGCAGCTCTTCTGAAATCCAAATATTTGGAGCACTTGAACAGCA 640
QY 494 CTGATCCCAATAGTGTCTTCCGATGTCAGATGTCAGAAATGGTTTGTATCATCATGGTC 553
Db 641 CTGACCCCAACAGTGTCTTCCGATGTCAGATGTCAGAAATGGTTTGTATCATCATGGTC 700
QY 554 ATGTTTGTATTGTGTTTGAACCTA CTGCGGACCTTAGTACCTTATGATTTTATTAAGAAACA 613
Db 701 ATGTTTGTATTGTGTTTGAACCTA CTGCGGACCTTAGTACCTTATGATTTTATTAAGAAACA 760
QY 614 GCTTTCTGCCATTTCCAAATTTGACCATGACGACATGCGGATGCGGATGCGGATGCGGATG 673
Db 761 GTTTTCTGCCATTTCCAAATTTGACCATGACGACATGCGGATGCGGATGCGGATGCGGATG 820
QY 674 TAAATTTTTTACATCATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 733
Db 821 TAAATTTTTTACATCATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 880
QY 734 TTGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 793
Db 881 TTGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 940
QY 794 TCAGAAACACAGATATCAAAAGTTGTTGACTTTGGAGTGCACGATGATGATGATGATGATG 853
Db 941 TGAAGAAACACAGATATCAAAAGTTGTTGACTTTGGAGTGCACGATGATGATGATGATG 1000
QY 854 ACAGTACTTTGCTGTACTACCGGACACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 913
Db 1001 ATAGTACTTTGCTGTACTACCGGACACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 1060
QY 914 GGTCTCAGCCTTTGATGATGTTTGGAGCATAGGTTGCTTCTTATTAATAATAATAATAATA 973
Db 1061 GGTCTCAGCCTTTGATGATGTTTGGAGCATAGGTTGCTTCTTATTAATAATAATAATAATA 1120
QY 974 TCACAGTCTTTGAGACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033
Db 1121 TCACAGTCTTTGAGACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1180
QY 1034 GACCCATACCAACACACATGATTTTCAAGAAACCAAGAAACCGCAAGTATTTTCCACCAATACC 1093
Db 1181 GACCCATACCAAGCAGCATATGATTCAGAGAGACCAAGAAACCGCAAGTATTTTCCACCAATACC 1240
QY 1094 AGCTAGATTGGGATGAACACAGTCTCTGCTGTAGATATGTTTAGGAGACGCTGCAACCGT 1153
Db 1241 AGCTAGATTGGGACGAGCATAGTTCAGCTGGAGATATGTTTAGGAGACGCTGCAACCGT 1300
QY 1154 TGAAGAAATTTATGCTTTGCTCATGATGAGAACATGAGAAACTGTTTGCCTGCTTTCGNA 1213
Db 1301 TAAAGGAATTTTATGCTGTGTCATGACGAAGAGCATGAGAAGCTGTTTGCCTGCTTTCGNA 1360
QY 1214 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGACAGCATCCTT 1273
Db 1361 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGACAGCATCCTT 1420
QY 1274 TCTTTGACTTATTAAGAAAGAAATGA 1299
Db 1421 TCTTTGACTTATTAAGAAAGAAATGA 1446
```

## RESULT 11

US-09-810-671-3

; Sequence 3, Application US/09810671

; Publication No. US20020076783A1

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; FILE REFERENCE: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE OF INVENTION: THEREOF

; FILE REFERENCE: CL000758

; CURRENT APPLICATION NUMBER: US/09/810,671

; CURRENT FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0









**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 08:20:29 ; Search time 8553.23 Seconds  
(without alignments)  
10767.492 Million cell updates/sec

Title: US-10-801-671-1\_COPY\_72\_2327

Perfect score: 2256

Sequence: 1 gacactatccagtcatta.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents\_NA\_Main:

1: /cgn2\_6/ptodata/1/pna/PCTUS1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/PCTUS2 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/PCTUS COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US06 COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US07 COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US081 COMB.seq.\*  
8: /cgn2\_6/ptodata/1/pna/US082 COMB.seq.\*  
9: /cgn2\_6/ptodata/1/pna/US083 COMB.seq.\*  
10: /cgn2\_6/ptodata/1/pna/US084 COMB.seq.\*  
11: /cgn2\_6/ptodata/1/pna/US085 COMB.seq.\*  
12: /cgn2\_6/ptodata/1/pna/US086 COMB.seq.\*  
13: /cgn2\_6/ptodata/1/pna/US087 COMB.seq.\*  
14: /cgn2\_6/ptodata/1/pna/US088 COMB.seq.\*  
15: /cgn2\_6/ptodata/1/pna/US089 COMB.seq.\*  
16: /cgn2\_6/ptodata/1/pna/US090 COMB.seq.\*  
17: /cgn2\_6/ptodata/1/pna/US091 COMB.seq.\*  
18: /cgn2\_6/ptodata/1/pna/US092 COMB.seq.\*  
19: /cgn2\_6/ptodata/1/pna/US092B COMB.seq.\*  
20: /cgn2\_6/ptodata/1/pna/US093A COMB.seq.\*  
21: /cgn2\_6/ptodata/1/pna/US093B COMB.seq.\*  
22: /cgn2\_6/ptodata/1/pna/US094 COMB.seq.\*  
23: /cgn2\_6/ptodata/1/pna/US095A COMB.seq.\*  
24: /cgn2\_6/ptodata/1/pna/US095B COMB.seq.\*  
25: /cgn2\_6/ptodata/1/pna/US095C COMB.seq.\*  
26: /cgn2\_6/ptodata/1/pna/US095D COMB.seq.\*  
27: /cgn2\_6/ptodata/1/pna/US096A COMB.seq.\*  
28: /cgn2\_6/ptodata/1/pna/US096B COMB.seq.\*  
29: /cgn2\_6/ptodata/1/pna/US096C COMB.seq.\*  
30: /cgn2\_6/ptodata/1/pna/US096D COMB.seq.\*  
31: /cgn2\_6/ptodata/1/pna/US096E COMB.seq.\*  
32: /cgn2\_6/ptodata/1/pna/US097A COMB.seq.\*  
33: /cgn2\_6/ptodata/1/pna/US097B COMB.seq.\*  
34: /cgn2\_6/ptodata/1/pna/US097C COMB.seq.\*  
35: /cgn2\_6/ptodata/1/pna/US098A COMB.seq.\*  
36: /cgn2\_6/ptodata/1/pna/US098B COMB.seq.\*  
37: /cgn2\_6/ptodata/1/pna/US098C COMB.seq.\*  
38: /cgn2\_6/ptodata/1/pna/US098D COMB.seq.\*  
39: /cgn2\_6/ptodata/1/pna/US099A COMB.seq.\*  
40: /cgn2\_6/ptodata/1/pna/US099B COMB.seq.\*  
41: /cgn2\_6/ptodata/1/pna/US099C COMB.seq.\*  
42: /cgn2\_6/ptodata/1/pna/US099D COMB.seq.\*  
43: /cgn2\_6/ptodata/1/pna/US099E COMB.seq.\*

44: /cgn2\_6/ptodata/1/pna/US099F COMB.seq.\*  
45: /cgn2\_6/ptodata/1/pna/US099G COMB.seq.\*  
46: /cgn2\_6/ptodata/1/pna/US100A COMB.seq.\*  
47: /cgn2\_6/ptodata/1/pna/US100B COMB.seq.\*  
48: /cgn2\_6/ptodata/1/pna/US101A COMB.seq.\*  
49: /cgn2\_6/ptodata/1/pna/US101B COMB.seq.\*  
50: /cgn2\_6/ptodata/1/pna/US102A COMB.seq.\*  
51: /cgn2\_6/ptodata/1/pna/US102B COMB.seq.\*  
52: /cgn2\_6/ptodata/1/pna/US103A COMB.seq.\*  
53: /cgn2\_6/ptodata/1/pna/US103B COMB.seq.\*  
54: /cgn2\_6/ptodata/1/pna/US104A COMB.seq.\*  
55: /cgn2\_6/ptodata/1/pna/US104B COMB.seq.\*  
56: /cgn2\_6/ptodata/1/pna/US105A COMB.seq.\*  
57: /cgn2\_6/ptodata/1/pna/US105B COMB.seq.\*  
58: /cgn2\_6/ptodata/1/pna/US106A COMB.seq.\*  
59: /cgn2\_6/ptodata/1/pna/US107A COMB.seq.\*  
60: /cgn2\_6/ptodata/1/pna/US107B COMB.seq.\*  
61: /cgn2\_6/ptodata/1/pna/US107C COMB.seq.\*  
62: /cgn2\_6/ptodata/1/pna/US107D COMB.seq.\*  
63: /cgn2\_6/ptodata/1/pna/US108A COMB.seq.\*  
64: /cgn2\_6/ptodata/1/pna/US108B COMB.seq.\*  
65: /cgn2\_6/ptodata/1/pna/US109A COMB.seq.\*  
66: /cgn2\_6/ptodata/1/pna/US109B COMB.seq.\*  
67: /cgn2\_6/ptodata/1/pna/US109C COMB.seq.\*  
68: /cgn2\_6/ptodata/1/pna/US110 COMB.seq.\*  
69: /cgn2\_6/ptodata/1/pna/US6000 COMB.seq.\*  
70: /cgn2\_6/ptodata/1/pna/US6001 COMB.seq.\*  
71: /cgn2\_6/ptodata/1/pna/US6002 COMB.seq.\*  
72: /cgn2\_6/ptodata/1/pna/US6003 COMB.seq.\*  
73: /cgn2\_6/ptodata/1/pna/US6004 COMB.seq.\*  
74: /cgn2\_6/ptodata/1/pna/US6005 COMB.seq.\*  
75: /cgn2\_6/ptodata/1/pna/US6006 COMB.seq.\*  
76: /cgn2\_6/ptodata/1/pna/US6007 COMB.seq.\*  
77: /cgn2\_6/ptodata/1/pna/US6008 COMB.seq.\*  
78: /cgn2\_6/ptodata/1/pna/US6009 COMB.seq.\*  
79: /cgn2\_6/ptodata/1/pna/US6010 COMB.seq.\*  
80: /cgn2\_6/ptodata/1/pna/US6011 COMB.seq.\*  
81: /cgn2\_6/ptodata/1/pna/US6012 COMB.seq.\*  
82: /cgn2\_6/ptodata/1/pna/US6013 COMB.seq.\*  
83: /cgn2\_6/ptodata/1/pna/US6014 COMB.seq.\*  
84: /cgn2\_6/ptodata/1/pna/US6015 COMB.seq.\*  
85: /cgn2\_6/ptodata/1/pna/US6016 COMB.seq.\*  
86: /cgn2\_6/ptodata/1/pna/US6017 COMB.seq.\*  
87: /cgn2\_6/ptodata/1/pna/US6018 COMB.seq.\*  
88: /cgn2\_6/ptodata/1/pna/US6019 COMB.seq.\*  
89: /cgn2\_6/ptodata/1/pna/US6020 COMB.seq.\*  
90: /cgn2\_6/ptodata/1/pna/US6021 COMB.seq.\*  
91: /cgn2\_6/ptodata/1/pna/US6022 COMB.seq.\*  
92: /cgn2\_6/ptodata/1/pna/US6023A COMB.seq.\*  
93: /cgn2\_6/ptodata/1/pna/US6023B COMB.seq.\*  
94: /cgn2\_6/ptodata/1/pna/US6024 COMB.seq.\*  
95: /cgn2\_6/ptodata/1/pna/US6025 COMB.seq.\*  
96: /cgn2\_6/ptodata/1/pna/US6026 COMB.seq.\*  
97: /cgn2\_6/ptodata/1/pna/US6027 COMB.seq.\*  
98: /cgn2\_6/ptodata/1/pna/US6028 COMB.seq.\*  
99: /cgn2\_6/ptodata/1/pna/US6029 COMB.seq.\*  
100: /cgn2\_6/ptodata/1/pna/US6030 COMB.seq.\*  
101: /cgn2\_6/ptodata/1/pna/US6031 COMB.seq.\*  
102: /cgn2\_6/ptodata/1/pna/US6032 COMB.seq.\*  
103: /cgn2\_6/ptodata/1/pna/US6033 COMB.seq.\*  
104: /cgn2\_6/ptodata/1/pna/US6034 COMB.seq.\*  
105: /cgn2\_6/ptodata/1/pna/US6035 COMB.seq.\*  
106: /cgn2\_6/ptodata/1/pna/US6036 COMB.seq.\*  
107: /cgn2\_6/ptodata/1/pna/US6037 COMB.seq.\*  
108: /cgn2\_6/ptodata/1/pna/US6038 COMB.seq.\*  
109: /cgn2\_6/ptodata/1/pna/US6039 COMB.seq.\*  
110: /cgn2\_6/ptodata/1/pna/US6040 COMB.seq.\*  
111: /cgn2\_6/ptodata/1/pna/US6041 COMB.seq.\*  
112: /cgn2\_6/ptodata/1/pna/US6042 COMB.seq.\*  
113: /cgn2\_6/ptodata/1/pna/US6043 COMB.seq.\*  
114: /cgn2\_6/ptodata/1/pna/US6044 COMB.seq.\*  
115: /cgn2\_6/ptodata/1/pna/US6045 COMB.seq.\*  
116: /cgn2\_6/ptodata/1/pna/US6046 COMB.seq.\*

117: /cgn2\_6/prodata/1/pna/us6047\_COMB.seq:  
118: /cgn2\_6/prodata/1/pna/us6048\_COMB.seq:  
119: /cgn2\_6/prodata/1/pna/us6049\_COMB.seq:  
120: /cgn2\_6/prodata/1/pna/us6050\_COMB.seq:  
121: /cgn2\_6/prodata/1/pna/us6051\_COMB.seq:  
122: /cgn2\_6/prodata/1/pna/us6052\_COMB.seq:  
123: /cgn2\_6/prodata/1/pna/us6053\_COMB.seq:  
124: /cgn2\_6/prodata/1/pna/us6054\_COMB.seq:  
125: /cgn2\_6/prodata/1/pna/us6055\_COMB.seq:  
126: /cgn2\_6/prodata/1/pna/us6056\_COMB.seq:  
127: /cgn2\_6/prodata/1/pna/us6057\_COMB.seq:  
128: /cgn2\_6/prodata/1/pna/us6058\_COMB.seq:  
129: /cgn2\_6/prodata/1/pna/us6059\_COMB.seq:

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Score | Length | DB ID | Description |
|------------|-------------|-------|--------|-------|-------------|
|------------|-------------|-------|--------|-------|-------------|

|    |        |       |       |     |                      |
|----|--------|-------|-------|-----|----------------------|
| 1  | 2256   | 100.0 | 2354  | 63  | US-10-801-671-1      |
| 2  | 2215.8 | 98.2  | 2497  | 48  | US-10-111-076A-1     |
| 3  | 2048   | 90.8  | 2488  | 44  | US-09-979-167-55     |
| 4  | 2047   | 90.7  | 2443  | 102 | US-60-324-185-20117  |
| 5  | 2047   | 90.7  | 2487  | 97  | US-60-278-258-3721   |
| 6  | 2046.6 | 90.7  | 2446  | 21  | US-09-332-978-9      |
| 7  | 1908.4 | 84.6  | 4035  | 54  | US-10-425-114-26852  |
| 8  | 1908.4 | 84.6  | 4035  | 54  | US-10-425-114A-26852 |
| 9  | 1825   | 80.9  | 2494  | 1   | PCT-US01-14827-2154  |
| 10 | 1825   | 80.9  | 2494  | 26  | US-09-577-408-3087   |
| 11 | 1635   | 72.5  | 1795  | 49  | US-10-170-335-20526  |
| 12 | 1634.6 | 72.5  | 1795  | 65  | US-10-932-349-580    |
| 13 | 1634.6 | 72.5  | 1795  | 120 | US-60-500-337-580    |
| 14 | 1632   | 72.3  | 1814  | 62  | US-10-756-149-49     |
| 15 | 1632   | 72.3  | 1814  | 117 | US-60-474-733-338    |
| 16 | 1300.6 | 57.7  | 1323  | 43  | US-09-959-164-7      |
| 17 | 1300.6 | 57.7  | 1323  | 43  | US-09-959-164A-7     |
| 18 | 1286   | 57.0  | 1446  | 51  | US-10-267-502-135    |
| 19 | 1263.2 | 56.0  | 1348  | 29  | US-09-652-128-8650   |
| 20 | 1263.2 | 56.0  | 1348  | 29  | US-09-652-917-2693   |
| 21 | 1263.2 | 56.0  | 1348  | 35  | US-09-813-206-755    |
| 22 | 1260   | 55.9  | 1261  | 90  | US-60-213-178-1312   |
| 23 | 1228.4 | 54.5  | 1247  | 22  | US-09-471-275-6193   |
| 24 | 1228.4 | 54.5  | 1247  | 22  | US-09-488-725B-5425  |
| 25 | 1228.4 | 54.5  | 1247  | 26  | US-09-552-317-5425   |
| 26 | 1185   | 52.5  | 1222  | 1   | PCT-US00-06765-17    |
| 27 | 1185   | 52.5  | 1222  | 1   | PCT-US00-06765-17    |
| 28 | 1185   | 52.5  | 1222  | 42  | US-09-950-083-657    |
| 29 | 1185   | 52.5  | 1222  | 42  | US-09-950-083B-657   |
| 30 | 1185   | 52.5  | 1222  | 48  | US-10-105-299-1157   |
| 31 | 1185   | 52.5  | 1222  | 64  | US-10-868-184-657    |
| 32 | 1185   | 52.5  | 1222  | 64  | US-10-868-184A-657   |
| 33 | 1185   | 52.5  | 1222  | 64  | US-10-868-184B-657   |
| 34 | 1185   | 52.5  | 1222  | 64  | US-10-868-194B-657   |
| 35 | 1185   | 52.5  | 1222  | 64  | US-10-896-164-657    |
| 36 | 1178.2 | 52.2  | 1456  | 58  | US-10-641-643-699    |
| 37 | 1099   | 48.7  | 1531  | 48  | US-10-144-771-2809   |
| 38 | 1099   | 48.7  | 1531  | 106 | US-60-360-207-2809   |
| 39 | 1097.4 | 48.6  | 1549  | 17  | US-09-137-248C-26    |
| 40 | 1097.4 | 48.6  | 1549  | 17  | US-09-137-248C-26    |
| 41 | 1097.4 | 48.6  | 1549  | 43  | US-09-959-164A-12    |
| 42 | 1097.4 | 48.6  | 1549  | 63  | US-10-825-177-26     |
| 43 | 1097.4 | 48.6  | 1549  | 117 | US-60-474-733-130    |
| 44 | 1097.2 | 48.6  | 1446  | 51  | US-10-267-502-138    |
| 45 | 1097   | 48.6  | 21234 | 63  | US-10-801-671-1      |

ALIGNMENTS

RESULT 1  
US-10-801-671-1  
; Sequence 1, Application US/10801671  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL000758DIV-III  
; CURRENT APPLICATION NUMBER: US/10/801,671  
; PRIOR FILING DATE: 2004-03-17  
; PRIOR APPLICATION NUMBER: 60/227,470  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 09/810,671  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2354  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-801-671-1

Query Match  
Best Local Similarity 100.0%; Score 2256; DB 63; Length 2354;  
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | GACACTCATCCAGTCATTATTTAGAACCAAGGTCCTTGAATGAGCGAGATTATTCGGGAC   | 60  |
| DB | 72  | GACACTCATCCAGTCATTATTTAGAACCAAGGTCCTTGAATGAGCGAGATTATTCGGGAC   | 131 |
| QY | 61  | CGGAGATACGTTGACGAATACAGGAATGACTTACTGTGAAGGATATGTTCTTAGACATTAT  | 120 |
| DB | 132 | CGGAGATACGTTGACGAATACAGGAATGACTTACTGTGAAGGATATGTTCTTAGACATTAT  | 191 |
| QY | 121 | CACGAGACATTTGAAAGCGGGTATCGAATTCACATGACGTAATCTTCACTCGGAGCAGG    | 180 |
| DB | 192 | CACGAGACATTTGAAAGCGGGTATCGAATTCACATGACGTAATCTTCACTCGGAGCAGG    | 251 |
| QY | 181 | AGAGGACGTCCTTAAAGGAAGCGCAATAGACACATGTTTCAAGTCATCACTGATTCGAAG   | 240 |
| DB | 252 | AGAGGACGTCCTTAAAGGAAGCGCAATAGACACATGTTTCAAGTCATCACTGATTCGAAG   | 311 |
| QY | 241 | AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACTGATCTGT      | 300 |
| DB | 312 | AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACTGATCTGT      | 371 |
| QY | 301 | CAAGTGGAGACGTTCTTAAGACGAAGATATCAAAATCGTGGACACTTTGGGTGAAGGCC    | 360 |
| DB | 372 | CAAGTGGAGACGTTCTTAAGACGAAGATATCAAAATCGTGGACACTTTGGGTGAAGGCC    | 431 |
| QY | 361 | TTTGGCAAGTTGTAGAGTCATTGATCATGGCATGGATGGCATGTCATGTAGCAGTGAAG    | 420 |
| DB | 432 | TTTGGCAAGTTGTAGAGTCATTGATCATGGCATGGATGGCATGTCATGTAGCAGTGAAG    | 491 |
| QY | 421 | ATCGTAAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAG    | 480 |
| DB | 492 | ATCGTAAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAG    | 551 |
| QY | 481 | CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGATGTTT      | 540 |
| DB | 552 | CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGATGTTT      | 611 |
| QY | 541 | GATCATCATGTCATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT     | 600 |
| DB | 612 | GATCATCATGTCATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT        | 671 |
| QY | 601 | ATTAAGAAAAACAGCTTTCTCCATTTCCTCAATTCGAAATGACCAATCAGCAGATGCGGTAT | 660 |
| DB | 672 | ATTAAGAAAAACAGCTTTCTCCATTTCCTCAATTCGAAATGACCAATCAGCAGATGCGGTAT | 731 |
| QY | 661 | ATCTGCCAGTCAATAAATTTTTTACATCATATTAATTAATTAATTAATTAATTAATTAAT   | 720 |



Db 732 ATCTCCAGTCAATAAATTTTATCATATAATAAATTAACCCATACAGATCTGAAGCT 791  
 QY 721 GAAATATTTTGTGTGAAGTCTGACTATGATGATCAAAATATAAATCTAAATGAAAGT 780  
 Db 792 GAAATATTTTGTGTGAAGTCTGACTATGATGATCAAAATATAAATCTAAATGAAAGT 851  
 QY 781 GATGAACGCACACTGAAACACACAGATATCAAAATGTTGACTTTGGAAGTGCACATAT 840  
 Db 852 GATGAACGCACACTGAAACACACAGATATCAAAATGTTGACTTTGGAAGTGCACATAT 911  
 QY 841 GATGATGAACATCAAGTACTTTGTGTCTACCCGCGCACTACAGAGCTCCCGAGGTCAAT 900  
 Db 912 GATGATGAACATCAAGTACTTTGTGTCTACCCGCGCACTACAGAGCTCCCGAGGTCAAT 971  
 QY 901 TTGGCTTTAGGTGGTCTCAGCTTTGTGATGTTTGGAGCATAGTGTGATTTTATTGAA 960  
 Db 972 TTGGCTTTAGGTGGTCTCAGCTTTGTGATGTTTGGAGCATAGTGTGATTTTATTGAA 1031  
 QY 961 TATTACCTTTGGTTTACAGTCTTTTACAGTCTATGATGATTAAGAGCACTGGCAATGATG 1020  
 Db 1032 TATTACCTTTGGTTTACAGTCTTTTACAGTCTATGATGATTAAGAGCACTGGCAATGATG 1091  
 QY 1021 GAAACGAATATTAGGACCCATACCAACACACATGATTTCAAGAAACCAAGAAACGCAAGTAT 1080  
 Db 1092 GHACGAATATTAGGACCCATACCAACACACATGATTTCAAGAAACCAAGAAACGCAAGTAT 1151  
 QY 1081 TTTTCAACATAACAGCTAGATTGGGATGAAACACAGTCTTGTCTGTAGATATGTTAGGAGA 1140  
 Db 1152 TTTTCAACATAACAGCTAGATTGGGATGAAACACAGTCTTGTCTGTAGATATGTTAGGAGA 1211  
 QY 1141 CGCTGCAACCGTTGAAGAAATTTATGCTTTGTCTGATGATGAAGAAACATGAGAAACCTGTTT 1200  
 Db 1212 CGCTGCAACCGTTGAAGAAATTTATGCTTTGTCTGATGATGAAGAAACATGAGAAACCTGTTT 1271  
 QY 1201 GACCTGGTTTGAAGAAATTTATGATGATCAACTCAAGAAATTAACCTTGGATGAAGCA 1260  
 Db 1272 GACCTGGTTTGAAGAAATTTATGATGATCAACTCAAGAAATTAACCTTGGATGAAGCA 1331  
 QY 1261 TTGACAGCTCTCTTTTGTGACTTTTAAAGAAATGAAGTGAAGTCAAGTGGTCTTAC 1320  
 Db 1332 TTGACAGCTCTCTTTTGTGACTTTTAAAGAAATGAAGTGAAGTCAAGTGGTCTTAC 1391  
 QY 1321 TATATACCTCTTGAAGAGATTAATTAAGACTGTGTCAGTCAACTAAACATCTTAATAT 1380  
 Db 1392 TATATACCTCTTGAAGAGATTAATTAAGACTGTGTCAGTCAACTAAACATCTTAATAT 1451  
 QY 1381 TTTTGTAAACATTAATTTTGTACAGTTAAAGTAAATTTGATTTGTTGATCAAA 1440  
 Db 1452 TTTTGTAAACATTAATTTTGTACAGTTAAAGTAAATTTGATTTGTTGATCAAA 1511  
 QY 1441 TAGCATTAATTAACCTTTTGAAGCAAGTATGCTTGTATGATGATTAAGAAATTAAGAAAT 1500  
 Db 1512 TAGCATTAATTAACCTTTTGAAGCAAGTATGCTTGTATGATGATTAAGAAATTAAGAAAT 1571  
 QY 1501 AATTTTCTTTTGAAGTAAATTAACCTTTTGAAGTAAATCTTTGATGATGATGATGATGAT 1560  
 Db 1572 AATTTTCTTTTGAAGTAAATTAACCTTTTGAAGTAAATCTTTGATGATGATGATGATGAT 1631  
 QY 1561 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 Db 1632 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1691  
 QY 1621 AGTAAAGGAATCTTGAAGTAAATTAACCTTTTGAAGTAAATCTTTTATATATATATATAT 1680  
 Db 1692 AGTAAAGGAATCTTGAAGTAAATTAACCTTTTGAAGTAAATCTTTTATATATATATATAT 1751  
 QY 1681 TTTTGAAGTAAATCTTTTGAAGTAAATTAACCTTTTGAAGTAAATCTTTTATATATATATAT 1740  
 Db 1752 TTTTGAAGTAAATCTTTTGAAGTAAATTAACCTTTTGAAGTAAATCTTTTATATATATATAT 1811  
 QY 1741 CTCTAGATGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
 Db 1812 CTCTAGATGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1871

QY 1801 TAAATTTTAAAGTGTGTATTTTCTTTTCAATGGGTGATGTCAAGGTGATAACAGACATTC 1860  
 Db 1872 TAAATTTTAAAGTGTGTATTTCTTTTCAATGGGTGATGTCAAGGTGATAACAGACATTC 1931  
 QY 1861 ATGGAAGGCATGCAAGTTTGTCCATTTGTCACAGTTTGTGTTTAAATAACACACATACACT 1920  
 Db 1932 ATGGAAGGCATGCAAGTTTGTCCATTTGTCACAGTTTGTGTTTAAATAACACACATACACT 1991  
 QY 1921 TTAATTAAGATTAATAATCTAACTGGAAGTCAAGTTTGGAAATGACATTTTCAAGTATG 1980  
 Db 1992 TTAATTAAGATTAATAATCTAACTGGAAGTCAAGTTTGGAAATGACATTTTCAAGTATG 2051  
 QY 1981 TTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
 Db 2052 TTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2111  
 QY 2041 AAGTCTTTAGGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTA 2100  
 Db 2112 AAGTCTTTAGGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTA 2171  
 QY 2101 ACCTATAAGAAATTAAGTTTATTAATTTAGGCAATTTATGTTCTGTGATAATTTCTTACGGGAG 2160  
 Db 2172 ACCTATAAGAAATTAAGTTTATTAATTTAGGCAATTTATGTTCTGTGATAATTTCTTACGGGAG 2231  
 QY 2161 AAAGAGGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTTCCAGAAATTT 2220  
 Db 2232 AAAGAGGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTTCCAGAAATTT 2291  
 QY 2221 AATTGATTTGGTTACATAAATCTTTTGGACTTTCAGAA 2256  
 Db 2292 AATTGATTTGGTTACATAAATCTTTTGGACTTTCAGAA 2327

RESULT 2

US-10-111-076A-1

; Sequence 1, Application US/10111076A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MAO, Yumin  
 ; APPLICANT: XIE, YI  
 ; TITLE OF INVENTION: A NEW POLYPEPTIDE - NEW CELL CYCLE-REGULATING PROTEIN 53 AND A  
 ; FILE REFERENCE: 2310/51152  
 ; CURRENT APPLICATION NUMBER: US/10/111,076A  
 ; CURRENT FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: PCT/CN/00328  
 ; PRIOR FILING DATE: 2000-10-16  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 1

US-10-111-076A-1

; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURES:  
 ; NAME/KEY: CDS  
 ; LOCATION: (111)..(1556)  
 ; OTHER INFORMATION:  
 ; US-10-111-076A-1

Query Match 98.2%; Score 2215.8; DB 48; Length 2497;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 2220; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 14 GTCAATTTTAAAGCAAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 73  
 Db 271 GTCAATTTTAAAGCAAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 330  
 QY 74 ACGAATACAGGAATGACTACTGTGAAGTATGTTCTTAGACATTTATCACAGACATTTG 133  
 Db 331 ACGAATACAGGAATGACTACTGTGAAGTATGTTCTTAGACATTTATCACAGGCAATG 390  
 QY 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAAGTCCGAGCAGGAGAGCAATCTTA 193

Db 391 AAAGGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGCGAGCGAGGAGCGAGTCCTA 450  
Qy 194 AAAGGAAGCGCAATAGACACACTGTTCAAGTCATCAGTCACGTTTCGAAGAGCCACCGAAGGA 253  
Db 451 AAAGGAAGCGCAATAGACACACTGTTCAAGTCATCAGTCACGTTTCGAAGAGCCACCGAAGGA 510  
Qy 254 AAAGATCCAGGAGTATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGACG 313  
Db 511 AAAGATCCAGGAGTATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGACG 570  
Qy 314 TTCTAAGAGCAAGATATGAATCTGGACACTTTGGGTGAAGGAGCCCTTGGCAAGGTTG 373  
Db 571 TTCTAAGAGCAAGATATGAATCTGGACACTTTGGGTGAAGGAGCCCTTGGCAAGGTTG 630  
Qy 374 TAGAGTGCAATGATCATCGCATGGATGCGCATGTAGCTAGCAATCGTAAAGAAATG 433  
Db 631 TAGAGTGCAATGATCATCGCATGGATGCGCATGTAGCTAGCAATCGTAAAGAAATG 690  
Qy 434 TAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 493  
Db 691 TAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 750  
Qy 494 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGATGCTTGCATCATCATGTC 553  
Db 751 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGATGCTTGCATCATCATGTC 810  
Qy 554 ATGTTTGTATGTTGTTGAACTACTGCGACCTAGTACTTACGATTTCAATAAAGAAACA 613  
Db 811 ATGTTTGTATGTTGTTGAACTACTGCGACCTAGTACTTACGATTTCAATAAAGAAACA 870  
Qy 614 GCTTTCGCCATTTCAAATGACCAATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 930  
Db 871 GCTTTCGCCATTTCAAATGACCAATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 930  
Qy 674 TAAATTTTTACATCATTAATTAATTAACCCATACAGATCTGAAGCTCAAAATATTTGT 733  
Db 931 TAAATTTTTACATCATTAATTAATTAACCCATACAGATCTGAAGCTCAAAATATTTGT 990  
Qy 734 TTGTGAAGCTGACATGATGATCAATATAATCTTAAATGAACGTCATGAAGCGCAC 793  
Db 991 TTGTGAAGCTGACATGATGATCAATATAATCTTAAATGAACGTCATGAAGCGCAC 1050  
Qy 794 TGAATAACACAGATATCAAAGTTGTTGACTTTGGAAGTGCAACGTCATGATGAACATC 853  
Db 1051 TGAATAACACAGATATCAAAGTTGTTGACTTTGGAAGTGCAACGTCATGATGAACATC 1110  
Qy 854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGTCATTTGGCTTTAGGTT 913  
Db 1111 ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGTCATTTGGCTTTAGGTT 1170  
Qy 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTATTGAATATTACCTTGGTT 973  
Db 1171 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTATTGAATATTACCTTGGTT 1230  
Qy 974 TCACAGTCTTTCAGACTCATGATAGTAAGAGCACCTCGCAATGATGAACGAATATTAG 1033  
Db 1231 TCACAGTCTTTCAGACTCATGATAGTAAGAGCACCTCGCAATGATGAACGAATATTAG 1290  
Qy 1034 GACCATACACACACATGATTCAGAAAACAGAAAAAGCGAGTATTTCACCAATACC 1093  
Db 1291 GACCATACACACACATGATTCAGAAAACAGAAAAAGCGAGTATTTCACCAATACC 1350  
Qy 1094 AGCTAGATTGGATGAAACACAGTCTCTGTTGATGATATGTTAGGAGAGCGCTGCAACCGT 1153  
Db 1351 AGCTAGATTGGATGAAACACAGTCTCTGTTGATGATATGTTAGGAGAGCGCTGCAACCGT 1410  
Qy 1154 TGAAGGAATTTATGCTTTGTCTATGATGAAGAACATGAGAAATCTGTTGACCTGGTTCGAA 1213  
Db 1411 TGAAGGAATTTATGCTTTGTCTATGATGAAGAACATGAGAAATCTGTTGACCTGGTTCGAA 1470  
Qy 1214 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGAGCATCCTT 1273  
Db 1471 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGAGCATCCTT 1530

RESULT 3  
US-09-979-167-55  
; Sequence 55, Application US/09979167

Qy 1274 TCTTTGACTTATTAAAAAGAAATGGAATCAAGTGGTCTTACTATATCTCTCT 1333  
Db 1531 TCTTTGGCTTATTAAAAAGAAATGGAATCAAGTGGTCTTACTATATCTCTCT 1590  
Qy 1334 AGAAGAGTACTTAAAGACTGTGTCACTCACTCAATCAATCTTAATATTTTGTAAACATT 1393  
Db 1591 AGAAGAGTACTTAAAGACTGTGTCACTCACTCAATCAATCTTAATATTTTGTAAACATT 1650  
Qy 1394 AAAATTTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC 1453  
Db 1651 AAAATTTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC 1710  
Qy 1454 TTGTTAAGCAAGTATGTTCTTGAATGCAATGCAAAAAATAAATAATTTCTTTT 1513  
Db 1711 TTGTTAAGCAAGTATGTTCTTGAATGCAATGCAAAAAATAAATAATTTCTTTT 1770  
Qy 1514 GAAATTTACCATTTTAAATACCTTTGAAATATCTTTGTGTCAGTGATAATGTTG 1573  
Db 1771 GAAATTTACCATTTTAAATACCTTTGAAATATCTTTGTGTCAGTGATAATGTTG 1830  
Qy 1574 ATCTTGCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTGAGTAAAGGAAAT 1633  
Db 1831 ATCTTGCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTGAGTAAAGGAAAT 1890  
Qy 1634 CTTGACTACTTTTATATTCTTAAAGGAATATTTCTTTATATATCTTCAAACTTAAAC 1693  
Db 1891 CTTGACTACTTTTATATTCTTAAAGGAATATTTCTTTATATATCTTCAAACTTAAAC 1950  
Qy 1694 TTTTAAAGTTTTCTTCTGTAATTTGTTGAAGCGGTGATTTATTTAACTCTAGATAAGCA 1753  
Db 1951 TTTTAAAGTTTTCTTCTGTAATTTGTTGAAGCGGTGATTTATTTAACTCTAGATAAGCA 2010  
Qy 1754 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGATTTCTATTAAATTTTAAAGT 1813  
Db 2011 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGATTTCTATTAAATTTTAAAGT 2070  
Qy 1814 TTGTATTCTTTTTCATTGGGTGATGTCAGGGTGATAAACACACATTCATGGAAGGCAATG 1873  
Db 2071 TTGTATTCTTTTTCATTGGGTGATGTCAGGGTGATAAACACACATTCATGGAAGGCAATG 2130  
Qy 1874 CAGTTTGTCCATTTGTGACAGTTTCTTAAATAAACCACATACACATTTATTAAAGATTA 1933  
Db 2131 CAGTTTGTCCATTTGTGACAGTTTCTTAAATAAACCACATACACATTTATTAAAGATTA 2190  
Qy 1934 AAATCTAACTGAAAGTCAGCTTGGAAAAATGGAATTTCCAAATTTCCAAAGTATGTTGGTGAAGTCAC 1993  
Db 2191 AAATCTAACTGAAAGTCAGCTTGGAAAAATGGAATTTCCAAAGTATGTTGGTGAAGTCAC 2250  
Qy 1994 AGATATAAAAAATAGAAATTCATGATGAGAGGTTTCAGTTTAAATCCAAAGTCTTTAGGAG 2053  
Db 2251 AGATATAAAAAATAGAAATTCATGATGAGAGGTTTCAGTTTAAATCCAAAGTCTTTAGGAG 2310  
Qy 2054 TCTTAACTATGGCCAGCATCTGTTTATCAATAGCATAAATACCTTAAACCTTATAGAAAT 2113  
Db 2311 TCTTAACTATGGCCAGCATCTGTTTATCAATAGCATAAATACCTTAAACCTTATAGAAAT 2370  
Qy 2114 AAGTTTATTAATAGGCAATTTATGTCGTGATTAATCTTACGGGAGAGAAAGGATTTGA 2173  
Db 2371 AAGTTTATTAATAGGCAATTTATGTCGTGATTAATCTTACGGGAGAGAAAGGATTTGA 2430  
Qy 2174 TTGGAAGCAGTTTGGGAAGAAAGTGTCTGTAATTTCCAGAAATTTAAATTTGATTTGTTA 2233  
Db 2431 TTGGAAGCAGTTTGGGAAGAAAGTGTCTGTAATTTCCAGAAATTTAAATTTGATTTGTTA 2490  
Qy 2234 CATAAAC 2240  
Db 2491 CATAAAC 2497

```

; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; APPLICANT: SUDERSANAM, SUCHA
; TITLE OF INVENTION: PROTEIN KINASES
; FILE REFERENCE: 038602/1273
; CURRENT APPLICATION NUMBER: US/09/979,167
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (562)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-09-979-167-55

Query Match          90.8%; Score 2048; DB 44; Length 2488;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2148; Conservative 0; Mismatches 1; Indels 90; Gaps 1;

QY 14  GTCAATATTAGAACAAAGGTCCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 73
DB      |||
QY 74  ACGAATACAGGAATGACTACTGCTGAAGGATATGTTCTTAGACATTATCACAGACATTG 133
DB      |||
QY 398 ACGAATACAGGAATGACTACTGCTGAAGGATATGTTCTTAGACATTATCACAGACATTG 457
DB      |||
QY 134  AAAGCGGGTATCGAATCCACTCGAGTAATCTTCAGTCCGACGAGGAGGAGCAGTCCTA 193
DB      |||
QY 458  AAAGCGGGTATCGAATCCACTCGAGTAATCTTCAGTCCGACGAGGAGGAGCAGTCCTA 517
DB      |||
QY 194  AAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAAAGACCCACGAAGA 253
DB      |||
QY 518  AAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCG----- 561
DB      |||
QY 254  AAAGATCCAGGAGTATAGGAGTATGAGGAGGTCACCTGATCTGTCAAGTGAGAGAG 313
DB      |||
QY 562  ----- 561
DB      |||
QY 314  TTCTAAGACAGATATGAAATCGTGGACACTTTGSGTGAAGAGCGCTTTGGCAAGTTG 373
DB      |||
QY 562  -----NATGAAATCGTGGACACTTTGGTGAAGAGCGCTTTGGCAAGTTG 607
DB      |||
QY 374  TAGAGTGCATTGATCATGGCATGGATGGCATGCATGTAGCAGTGAATAATCGTAAAGAAATG 433
DB      |||
QY 608  TAGAGTGCATTGATCATGGCATGGATGGCATGCATGTAGCAGTGAATAATCGTAAAGAAATG 667
DB      |||
QY 434  TAGGCGGTTACCGTGAAGCAGTCGTTGAGAAATCCAGATATAGAGCACTTAAATAGTA 493
DB      |||
QY 668  TAGGCGGTTACCGTGAAGCAGTCGTTGAGAAATCCAGATATAGAGCACTTAAATAGTA 727
DB      |||
QY 494  CTGATCCCAATAGTCCTTCGATGTCAGATGTCAGATGTCAGATGTCATCATCATGTC 553
DB      |||
QY 728  CTGATCCCAATAGTCCTTCGATGTCAGATGTCAGATGTCAGATGTCATCATCATGTC 787
DB      |||
QY 554  ATGTTTGTATTGTTTGAACCTACTGGGACTTAGTACTTACCATTTCAATTAAGAAAAACA 613
DB      |||
QY 788  ATGTTTGTATTGTTTGAACCTACTGGGACTTAGTACTTACCATTTCAATTAAGAAAAACA 847
DB      |||
QY 614  GCTTTCTGCAATTCAAATGACCAATCAGGACAGATGCGGATCAGATCTGCCAGTCAA 673
DB      |||
QY 848  GCTTTCTGCAATTCAAATGACCAATCAGGACAGATGCGGATCAGATCTGCCAGTCAA 907
DB      |||
QY 674  TAAATTTTACATCATATAAATTAACCCATACAGATCTGAGGCTGAAATATTTTGT 733
DB      |||
QY 908  TAAATTTTACATCATATAAATTAACCCATACAGATCTGAGGCTGAAATATTTTGT 967
DB      |||

734  TTGTGAAGTCTGACATATGTAGTCAAAATATAATTTCTAAAAATGAAACGTGATGAACGCAC 793
DB      |||
968  TTGTGAAGTCTGACATATGTAGTCAAAATATAATTTCTAAAAATGAAACGTGATGAACGCAC 1027
QY      |||
794  TGAATAACACAGATATCAAAAGTTGTGACTTTTGGAGTGCACACGTATGATGATGAACATC 853
DB      |||
1028  TGAATAACACAGATATCAAAAGTTGTGACTTTTGGAGTGCACACGTATGATGATGAACATC 1087
QY      |||
854  ACAGTACTTTTGGTGTCTATCCCGGCACACTACAGAGCTCCCGAGGTCAATTTTGGGCTTTAG 913
DB      |||
1088  ACAGTACTTTTGGTGTCTATCCCGGCACACTACAGAGCTCCCGAGGTCAATTTTGGGCTTTAG 1147
QY      |||
914  GGTCTCAGCCTTTGATGATTTTGGAGCATAGTTGTCATTTCTATTGAATATTAACCTTGGTT 973
DB      |||
1148  GGTCTCAGCCTTTGATGATTTTGGAGCATAGTTGTCATTTCTATTGAATATTAACCTTGGTT 1207
QY      |||
974  TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGCAATGATGAACCAATATTAG 1033
DB      |||
1208  TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGCAATGATGAACCAATATTAG 1267
QY      |||
1034  GACCCATACCAACACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCAACATAACC 1093
DB      |||
1268  GACCCATACCAACACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCAACATAACC 1327
QY      |||
1094  AGCTAGATTGGATGAACACACAGTTCCTGCTGCTAGATATGTTAGGAGACGCTGCCAACCGT 1153
DB      |||
1328  AGCTAGATTGGATGAACACACAGTTCCTGCTGCTAGATATGTTAGGAGACGCTGCCAACCGT 1387
QY      |||
1154  TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACCATGAGAACCTGTTGACCTGGTTC 1213
DB      |||
1388  TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACCATGAGAACCTGTTGACCTGGTTC 1447
QY      |||
1214  GAACTGTAGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCAGAGCATCCTT 1273
DB      |||
1448  GAACTGTAGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCAGAGCATCCTT 1507
QY      |||
1274  TCTTTGACTTATTTAAAAAGAAATGAAATGGAATCAGTGGTCTTACTATATATCTTCTCT 1333
DB      |||
1508  TCTTTGACTTATTTAAAAAGAAATGAAATGGAATCAGTGGTCTTACTATATATCTTCTCT 1567
QY      |||
1334  AGAAGAGATTACTTAAGACTGTGTCAGTCACTCACTAAACATTCCTAATATTTTGTGAACATT 1393
DB      |||
1568  AGAAGAGATTACTTAAGACTGTGTCAGTCACTCACTAAACATTCCTAATATTTTGTGAACATT 1627
QY      |||
1394  AAATTTATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATAATTAAC 1453
DB      |||
1628  AAATTTATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATAATTAAC 1687
QY      |||
1454  TTGTTAAGCAAGTATGGTCTTTGATAATGCATTTAGAAAAATTTAAATTTTCTTTT 1513
DB      |||
1688  TTGTTAAGCAAGTATGGTCTTTGATAATGCATTTAGAAAAATTTAAATTTTCTTTT 1747
QY      |||
1514  GAAATACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATATGATG 1573
DB      |||
1748  GAAATACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATATGATG 1807
QY      |||
1574  ATCTTGCTTTTGTACATGGAGGTCACTCTGGAAGTGAATTTTGTAGTAAAGGAAAT 1633
DB      |||
1808  ATCTTGCTTTTGTACATGGAGGTCACTCTGGAAGTGAATTTTGTAGTAAAGGAAAT 1867
QY      |||
1634  CTTGACTACTTTTATATTTCTTAAAGGAATTTCTTTTATATATCTTCAAAATTTAGAACTTAAC 1693
DB      |||
1868  CTTGACTACTTTTATATTTCTTAAAGGAATTTCTTTTATATATCTTCAAAATTTAGAACTTAAC 1927
QY      |||
1694  TTTAAAGTCTTTCTTCTGTAATTTGTTGAAACGGGTATTTATTTAACTCTAGATAAGCA 1753
DB      |||
1928  TTTAAAGTCTTTCTTCTGTAATTTGTTGAAACGGGTATTTATTTAACTCTAGATAAGCA 1987
QY      |||
1754  GGTACTAGAAACCAAACTCAGAAAAATCTTTACTGTTAGAAATTTCTATTAAATTTTAAAGT 1813
DB      |||
1988  GGTACTAGAAACCAAACTCAGAAAAATCTTTACTGTTAGAAATTTCTATTAAATTTTAAAGT 2047
QY      |||
1814  TTGTATTTCTTTTCTATTGGGTGATGTCAGGGGTGATAACACAGACATTCATGGAAGGCATG 1873

```



```
Db 1463 TCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATACTTCTCT 1522
Qy 1334 AGAAGAGATTACTTAAGACTGTGTGAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1393
Db 1523 AGAAGAGATTACTTAAGACTGTGTGAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1582
Qy 1394 AAATTTATTTGTACAGTTAAAGTGTAATATTTGTATGTTTGTATCAATAGCATAATTAAC 1453
Db 1583 AAATTTATTTGTACAGTTAAAGTGTAATATTTGTATGTTTGTATCAATAGCATAATTAAC 1642
Qy 1454 TTGTTAAGCAAGTATGTTCTTGATAATGATTAAGAAATTAAGAAATTAATTTCTTTT 1513
Db 1643 TTGTTAAGCAAGTATGTTCTTGATAATGATTAAGAAATTAAGAAATTAATTTCTTTT 1702
Qy 1514 GAAATTTACATTTTAAATACCTTTTGAATATCTTTTGTGTCAGATGATAATGATGATG 1573
Db 1703 GAAATTTACATTTTAAATACCTTTTGAATATCTTTTGTGTCAGATGATAATGATGATG 1762
Qy 1574 ATCTTGCCCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTAGTAAAGGAAT 1633
Db 1763 ATCTTGCCCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTAGTAAAGGAAT 1822
Qy 1634 CTGACTACTTTTATATCTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAAC 1693
Db 1823 CTGACTACTTTTATATCTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAAC 1882
Qy 1694 TTTAAAGTTTTTCTCTGTAAATCTTGAAGGATGATTTATTTAACTAGATAAGCA 1753
Db 1883 TTTAAAGTTTTTCTCTGTAAATCTTGAAGGATGATTTATTTAACTAGATAAGCA 1942
Qy 1754 GGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTAATTAATTTTAAAGT 1813
Db 1943 GGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTAATTAATTTTAAAGT 2002
Qy 1814 TTGTAATCTTTTCAATTCAGTGTGATCTCAGGTTGATAACAGACATTTCTATGGAAGGCATG 1873
Db 2002 TTGTAATCTTTTCAATTCAGTGTGATCTCAGGTTGATAACAGACATTTCTATGGAAGGCATG 2062
Qy 1874 CAGTTTGTCCATGTGACAGTTTGTTTAATAAACACATACACACATTTTATTAAGATTA 1933
Db 2063 CAGTTTGTCCATGTGACAGTTTGTTTAATAAACACATACACACATTTTATTAAGATTA 2122
Qy 1934 AAATCTAACTGGAAGTCACTTGAAATGGAATGGAATTTCCAGATATGTTGGTGAAGTCA 1993
Db 2123 AAATCTAACTGGAAGTCACTTGGAATGGAATGGAATTTCCAGATATGTTGGTGAAGTCA 2182
Qy 1994 AGATATAAAATAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACCAAGTCTTTAGAG 2053
Db 2183 AGATATAAAATAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACCAAGTCTTTAGAG 2242
Qy 2054 TCTTAAACTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAAGAAAT 2113
Db 2243 TCTTAAACTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAAGAAAT 2302
Qy 2114 AAGTTTATTAATAGGCAATTTATGTTCTGTGATATTTCTTACGGGAGAAAGAGGATTTGA 2173
Db 2303 AAGTTTATTAATAGGCAATTTATGTTCTGTGATATTTCTTACGGGAGAAAGAGGATTTGA 2362
Qy 2174 TTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTCCAGATTTAATTTGATTTGGTTA 2233
Db 2363 TTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTCCAGATTTAATTTGATTTGGTTA 2422
Qy 2234 CATAAACTTTTTTGACTTCA 2252
Db 2423 CATAAACTTTTTTGACTTCA 2441
```

## RESULT 5

US-60-278-3721

; Sequence 3721, Application US/60278258

; GENERAL INFORMATION:

; APPLICANT: Morris, MacDonald

; APPLICANT: Lal, Preeti

```
; APPLICANT: Diap, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0010-1_P
; CURRENT APPLICATION NUMBER: US/60/278,258
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 17730
; SOFTWARE: PERL Program
; SEQ ID NO 3721
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 253847.3
US-60-278-3721
```

```
Query Match 90.7%; Score 2047; DB 97; Length 2487;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2148; Conservative 0; Mismatches 0; Indels 91; Gaps 1;

Qy 14 GTCATTATTTTAAAGCAAGGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 73
Db 338 GTCATTATTTTAAAGCAAGGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 397

Qy 74 ACGAATACAGGAATGACTACTGTGAAAGGATATGTTCTTAGACATTATCACAGACATTG 133
Db 398 ACGAATACAGGAATGACTACTGTGAAAGGATATGTTCTTAGACATTATCACAGACATTG 457

Qy 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGAGAGAGAGTCTCTA 193
Db 458 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGAGAGAGAGTCTCTA 517

Qy 194 AAAGGAGCGCAATAGACACATGTTCAAGTCATCAGTCACTGCGAAGAGCCACCGAAG 253
Db 518 AAAGGAGCGCAATAGACACATGTTCAAGTCATCAGTCACTGCGAAGAGCCACCGAAG 561

Qy 254 AAAGTCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGAG 313
Db 562 -----

Qy 314 TTCTAAGAGCAAGATATGAATCTGTGGACACTTTGGGTGAAGAGGCCCTTTGGCAAAGTTG 373
Db 562 -----ATGAATCTGTGGACACTTTGGGTGAAGAGGCCCTTTGGCAAAGTTG 606

Qy 374 TAGAGTGCATTTGATCATGGCATGGATGCGATGTCATGTAGCAGTGAATTCGTAAAAATG 433
Db 607 TAGAGTGCATTTGATCATGGCATGGATGCGATGTCATGTAGCAGTGAATTCGTAAAAATG 666

Qy 434 TAGGCCGTTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCACCTTAAATAGTA 493
Db 667 TAGGCCGTTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCACCTTAAATAGTA 726

Qy 494 CTGATCCCAATAGTGTCTTCCGATGTCGAGATGTCAGATGCTAGAAATGTTGATCATGTC 553
Db 727 CTGATCCCAATAGTGTCTTCCGATGTCGAGATGTCAGATGCTAGAAATGTTGATCATGTC 786

Qy 554 ATGTTTGTATTTGTTTGAACACTAGTGGACCTTAGTACTTACGATTTCAATTAAGAAACA 613
Db 787 ATGTTTGTATTTGTTTGAACACTAGTGGACCTTAGTACTTACGATTTCAATTAAGAAACA 846

Qy 614 GCTTTCTGCCATTTTCAAATTTGACCAATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 673
Db 847 GCTTTCTGCCATTTTCAAATTTGACCAATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 906

Qy 674 TAAATTTTATACATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 733
Db 907 TAAATTTTATACATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 966

Qy 734 TTGTAAGTCTGACTATGTAATTAATAATAATAATAATAATAATAATAATAATAATAATA 793
Db 967 TTGTAAGTCTGACTATGTAATTAATAATAATAATAATAATAATAATAATAATAATAATA 1026
```



STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: NEUTRUMT01  
CLONE: 33963  
SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
US-09-392-978-9

Query Match 90.7%; Score 2046.6; DB 21; Length 2446;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 2149; Conservative 0; Mismatches 4; Indels 90; Gaps 1;

|    |      |  |      |
|----|------|--|------|
| QY | 14   | GTCAATATTAGAACCAAGGTCCTTGAATGACGAGATATTATCGGACCGGAGATACGTTG    | 73   |
| DB | 288  | GTCAATATTAGAACCAAGGTCCTTGAATGACGAGATATTATCGGACCGGAGATACGTTG    | 347  |
| QY | 74   | ACGAATACAGGAATGACTACTCTGGAAGGATATGTTCTTAGACATTATCACAGACATTG    | 133  |
| DB | 348  | ACGAATACAGGANUGACTACTCTGGAAGGATATGTTCTTAGACATTATCACAGACATTG    | 407  |
| QY | 134  | AAAGCGGGTATCGAATCCACTGACGTAAATCTTCAGTCCGACGAGGAGAGACGTCTTA     | 193  |
| DB | 408  | AAAGCGGGTATCGAATCCACTGACGTAAATCTTCAGTCCGACGAGGAGAGACGTCTTA     | 467  |
| QY | 194  | AAAGGAAGCGGATAGACACTGTTCAAGTCATCAGTTCGAAAGACCCACGAAGA          | 253  |
| DB | 468  | AAAGGAAGCGGATAGACACTGTTCAAGTCATCAGTTCGAAAGACCCACGAAGA          | 515  |
| QY | 254  | AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTGTCTCAAAGTGGAGAG         | 313  |
| DB | 516  | -----  | 515  |
| QY | 314  | TTCTAAGAGCAAGATATGAATCGTGGACACTTTTGGGTGAAGGAGCCTTTGGCAAGTTG    | 373  |
| DB | 516  | -----AATCGGTGGACACTTTTGGGTGAAGGAGCCTTTGGCAAGTTG                | 557  |
| QY | 374  | TAGAGTGCAATGATCATGGCATGGATGGATGATGATGATGATGATGATGATGATGATG     | 433  |
| DB | 558  | TAGAGTGCAATGATCATGGCATGGATGGATGATGATGATGATGATGATGATGATGATG     | 617  |
| QY | 434  | TAGGCGGTACCGTGAAGCGTCTTCCAGATGCTAGAGTCCAGATGATGATGATGATGATG    | 493  |
| DB | 618  | TAGGCGGTACCGTGAAGCGTCTTCCAGATGCTAGAGTCCAGATGATGATGATGATGATG    | 677  |
| QY | 494  | CTGATCCCAATAGTGTCTTCCAGATGCTAGAGTCCAGATGATGATGATGATGATGATG     | 553  |
| DB | 678  | CTGATCCCAATAGTGTCTTCCAGATGCTAGAGTCCAGATGATGATGATGATGATGATG     | 737  |
| QY | 554  | ATGTTTGTATTGTTGAACCTACTGGGACTTACTGATGATGATGATGATGATGATGATGATG  | 613  |
| DB | 738  | ATGTTTGTATTGTTGAACCTACTGGGACTTACTGATGATGATGATGATGATGATGATGATG  | 797  |
| QY | 614  | GCTTCTGCCATTTCAAAATGACCATCAGGACATGCGGATGATGATGATGATGATGATGATG  | 673  |
| DB | 798  | GCTTCTGCCATTTCAAAATGACCATCAGGACATGCGGATGATGATGATGATGATGATGATG  | 857  |
| QY | 674  | TAAATTTTTCATCATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT   | 733  |
| DB | 858  | TAAATTTTTCATCATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT   | 917  |
| QY | 734  | TGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG    | 793  |
| DB | 918  | TGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG    | 977  |
| QY | 794  | TGAAAAACACAGATATCAAAAGTTGTTGATTTTGGAAAGTGCACGATGATGATGATGATG   | 853  |
| DB | 978  | TGAAAAACACAGATATCAAAAGTTGTTGATTTTGGAAAGTGCACGATGATGATGATGATG   | 1037 |
| QY | 854  | ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT   | 913  |
| DB | 1038 | ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT   | 1097 |
| QY | 914  | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATTTCTTATTGAATATTACCTTGGTT | 973  |

|    |      |  |      |
|----|------|--|------|
| DB | 1098 | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATTTCTTATTGAATATTACCTTGGTT | 1157 |
| QY | 974  | TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTCGCAATGATGAAACCAATATTAG   | 1033 |
| DB | 1158 | TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTCGCAATGATGAAACCAATATTAG   | 1217 |
| QY | 1034 | GACCCATACCAACACACATGATTCAGAAAAACGAAAAACGAAAGTATTTCCTACCAATACC  | 1093 |
| DB | 1218 | GACCCATACCAACACACATGATTCAGAAAAACGAAAAACGAAAGTATTTCCTACCAATACC  | 1277 |
| QY | 1094 | AGCTAGATTGGATGAAACACACAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG   | 1153 |
| DB | 1278 | AGCTAGATTGGATGAAACACACAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG   | 1337 |
| QY | 1154 | TGAAGGAATTTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG    | 1213 |
| DB | 1338 | TGAAGGAATTTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG    | 1397 |
| QY | 1214 | GAATGTTAGATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCATTCGAGCATCCTT  | 1273 |
| DB | 1398 | GAATGTTAGATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCATTCGAGCATCCTT  | 1457 |
| QY | 1274 | TCTTTGACTTATTAAGAAAGAAATGAAATGGAATGGAATGGAATGGAATGGAATGGAATG   | 1333 |
| DB | 1458 | TCTTTGACTTATTAAGAAAGAAATGAAATGGAATGGAATGGAATGGAATGGAATGGAATG   | 1517 |
| QY | 1334 | AGAAGAGATTAATTAAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG   | 1393 |
| DB | 1518 | AGAAGAGATTAATTAAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG   | 1577 |
| QY | 1394 | AAATTAATTTTGTACAGTTAAGTGAATTAATTTGTATGTTTGTATCAATAGCATATAATAC  | 1453 |
| DB | 1578 | AAATTAATTTTGTACAGTTAAGTGAATTAATTTGTATGTTTGTATCAATAGCATATAATAC  | 1637 |
| QY | 1454 | TTGTTAAGCAAGTATGCTTGTGATTAATGCAATAGAAAAATTAATAATTAATTTCTTTT    | 1513 |
| DB | 1638 | TTGTTAAGCAAGTATGCTTGTGATTAATGCAATAGAAAAATTAATAATTAATTTCTTTT    | 1697 |
| QY | 1514 | GAATTAATTTTAAATACCTTTGAAATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG   | 1573 |
| DB | 1698 | GAATTAATTTTAAATACCTTTGAAATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG   | 1757 |
| QY | 1574 | ATCTTGCTTTTGTACAGGAGTCACTCTGCAAGTGAATTTTGTAGTAAAGGAAAT         | 1633 |
| DB | 1758 | ATCTTGCTTTTGTACAGGAGTCACTCTGCAAGTGAATTTTGTAGTAAAGGAAAT         | 1817 |
| QY | 1634 | CTTGACTTTTATTAATTTAAAGGAAATTTCTTTATATATCTTCAAAATTTAGAACTTAAC   | 1693 |
| DB | 1818 | CTTGACTTTTATTAATTTAAAGGAAATTTCTTTATATATCTTCAAAATTTAGAACTTAAC   | 1877 |
| QY | 1694 | TTTAAAGTCTTCTGCTGATTTGCTGAAACGGGTGATTTATTAATTAATTAATTAATTAAT   | 1753 |
| DB | 1878 | TTTAAAGTCTTCTGCTGATTTGCTGAAACGGGTGATTTATTAATTAATTAATTAATTAAT   | 1937 |
| QY | 1754 | GCTACTAGAAACCAAACTCAGAAAAATGTTTACTGCTAGAAATTTTAAATTTTAAAGTG    | 1813 |
| DB | 1938 | GCTACTAGAAACCAAACTCAGAAAAATGTTTACTGCTAGAAATTTTAAATTTTAAAGTG    | 1997 |
| QY | 1814 | TGTATTTCTTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG    | 1873 |
| DB | 1998 | TGTATTTCTTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG    | 2057 |
| QY | 1874 | CAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG    | 1933 |
| DB | 2058 | CAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG    | 2117 |
| QY | 1934 | AAATCTAACTGAAAGTCAAGCTTGGAAAAACGCAATTTCCAAAGTATGTTTGGTGGTAC    | 1993 |
| DB | 2118 | AAATCTAACTGAAAGTCAAGCTTGGAAAAACGCAATTTCCAAAGTATGTTTGGTGGTAC    | 2177 |
| QY | 1994 | AGATATAAAAAATAGAAATTTCTGATGAGAGGTTTCAAGTTTAAATACCAAGTCTTTAGGAG | 2053 |



```

Db 2178 AGATATAAATAAGAAATCTGATGAGAGGTTTCAGTTTTTAATACCAAGTCCTTAGGAG 2237
Qy 2054 TCTTAACATTCGCCAGCATCTGTTTATCAATAGACATATAACGTAAACCTTATAGAAATT 2113
Db 2238 TCTTAACATTCGCCAGCATCTGTTTATCAATAGACATATAACGTAAACCTTATAGAAATT 2297
Qy 2114 AAGTTTATTAATAGGCAATTTATGTCGTGATAATCTTACGGGAGAAAGAGGATTGGA 2173
Db 2298 AAGTTTATTAATAGGCAATTTATGTCGTGATAATCTTACGGGAGAAAGAGGATTGGA 2357
Qy 2174 TTGGAACAGTTTGGGAAGAAAGTCGCTGCTGAAATTTCCAGAAATTTAATGATGGTTA 2233
Db 2358 TTGGAACAGTTTGGGAAGAAAGTCGCTGCTGAAATTTCCAGAAATTTAATGATGGTTA 2417
Qy 2234 CATAACTTTTTTGACTTCAGAAA 2256
Db 2418 CATAACTTTTTTGACTTCAAAAA 2440

RESULT 7
US-10-425-114-26852
; Sequence 26852, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26852
; LENGTH: 4035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4654-056-G4_FLI
US-10-425-114-26852

Query Match 84.6%; Score 1908.4; DB 54; Length 4035;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 327 ATATGAATTCGTGGACACTTTGGGTGAAGAGGACCTTTGGCAGAAAGTTGTAGAGTGCAATTGA 386
Db 2126 AGATGAATTCGTGGACACTTTGGGTGAAGAGGACCTTTGGCAGAAAGTTGTAGAGTGCAATTGA 2185
Qy 387 TCATGGCATGGATGGCATGGATAGCAGTGAATATCGTAAATCGTAAATATGAGGCGTTACCG 446
Db 2186 TCATGGCATGGATGGCATGGATAGCAGTGAATATCGTAAATCGTAAATATGAGGCGTTACCG 2245
Qy 447 TGAACAGCTCGTTCAGAAATCCAGATATTAGACACTTAAATAGTACTGATCCCAATAG 506
Db 2246 TGAACAGCTCGTTCAGAAATCCAGATATTAGACACTTAAATAGTACTGATCCCAATAG 2305
Qy 507 TGTCTTCGATGTGCCAGATGCTAGATGGTTTGATCATCATGATGGTCAATGTTGATTTGT 566
Db 2306 TGTCTTCGATGTGCCAGATGCTAGATGGTTTGATCATCATGATGGTCAATGTTGATTTGT 2365
Qy 567 GTTTGAATCTAGTGGACCTTAGTACTAGTATTCATTAAGAAACAGCTTCTGCCAATT 626
Db 2366 GTTTGAATCTAGTGGACCTTAGTACTAGTATTCATTAAGAAACAGCTTCTGCCAATT 2425
Qy 627 TCAAAATGACCACATCAGGAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA 686
Db 2426 TCAAAATGACCACATCAGGAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA 2485
Qy 687 TCATTAATTAATTAACCCATACAGATCTGAGCCTGAAATATTTTGTGTGAGAGTCGTA 746

```

```

Db 2486 TCATAATAAATAAACCCATACAGATCTGAAGCCTGAAATAATTTTTTGTGTAAGTCGTA 2545
Qy 747 CTATGTAGTCAAAATATAATTTCTAAATGAAACGTGATGAACGCACACTGAAACACACAGA 806
Db 2546 CTATGTAGTCAAAATATAATTTCTAAATGAAACGTGATGAACGCACACTGAAACACACAGA 2605
Qy 807 TATCAAAAGTTGTTGCACTTTGGAAAGTGCAACGATATGATGATGAACATCAACAGTACTTTGGT 866
Db 2606 TATCAAAAGTTGTTGCACTTTGGAAAGTGCAACGATATGATGATGAACATCAACAGTACTTTGGT 2665
Qy 867 GTCTACCCGGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCCTTG 926
Db 2666 GTCTACCCGGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCCTTG 2725
Qy 927 TGATGTTTGGAGCATAGTGTGCAATTTTATTAATTAATTTACCTTGGTTTTCACAGTCTTTTCA 986
Db 2726 TGATGTTTGGAGCATAGTGTGCAATTTTATTAATTAATTTACCTTGGTTTTCACAGTCTTTTCA 2785
Qy 987 GACTCATGATAGTAAAGAGCACCTGGCAATGATGAAGACGAATATAGGACCCATACCAACA 1046
Db 2786 GACTCATGATAGTAAAGAGCACCTGGCAATGATGAAGACGAATATAGGACCCATACCAACA 2845
Qy 1047 ACACATGATTCAGAAACCAAGAAACCGNAGTATTTTACCATTAACAGCTAGATTGGGA 1106
Db 2846 ACACATGATTCAGAAACCAAGAAACCGNAGTATTTTACCATTAACAGCTAGATTGGGA 2905
Qy 1107 TGAACACAGTTCTGCTGCTAGATATGTTAGAGACGCTGCAACCGTTGAAAGGAATTTAT 1166
Db 2906 TGAACACAGTTCTGCTGCTAGATATGTTAGAGACGCTGCAACCGTTGAAAGGAATTTAT 2965
Qy 1167 GCTTTGTCATGATGAAGAACATGAGAACTGTTTGGCTGTTGCTGTTGCAAGAAATTTAGATA 1226
Db 2966 GCTTTGTCATGATGAAGAACATGAGAACTGTTTGGCTGTTGCTGTTGCAAGAAATTTAGATA 3025
Qy 1227 TGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGAGCATCTTCTTTTGACTTATT 1286
Db 3026 TGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGAGCATCTTCTTTTGACTTATT 3085
Qy 1287 AAAAAAGAAATGAAATGGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGATTAAT 1346
Db 3086 AAAAAAGAAATGAAATGGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGATTAAT 3145
Qy 1347 TAAGACTGTGTGCTCACTCAATCAATTTCTTAATATTTTGTGTAACATTAATTTTGTGA 1406
Db 3146 TAAGACTGTGTGCTCACTCAATCAATTTCTTAATATTTTGTGTAACATTAATTTTGTGA 3205
Qy 1407 CAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATAAATTAACCTGTTTAAGCAAGT 1466
Db 3206 CAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATAAATTAACCTGTTTAAGCAAGT 3265
Qy 1467 ATGCTCTTGAATAATGCAATAGAAAAATTAATAATTTTCTTTTGTGAAATTTACCAATT 1526
Db 3266 ATGCTCTTGAATAATGCAATAGAAAAATTAATAATTTTCTTTTGTGAAATTTACCAATT 3325
Qy 1527 TTAATAACCTTTGAAATATCTTTGTGTCAGTGATTAATGATGATCTTGTGCTTTTG 1586
Db 3326 TTAATAACCTTTGAAATATCTTTGTGTCAGTGATTAATGATGATCTTGTGCTTTTG 3385
Qy 1587 TACATGAGGTCACCTCTGAAAGTGAATTTTTCAGTAAAGGAAATCTTCACTACTTTTA 1646
Db 3386 TACATGAGGTCACCTCTGAAAGTGAATTTTTCAGTAAAGGAAATCTTCACTACTTTTA 3445
Qy 1647 TATTCTTAAAGGAATTTCTTTTATATATCTTCAAAATTTAGAACTTTAACTTTAAAGTTT 1706
Db 3446 TATTCTTAAAGGAATTTCTTTTATATATCTTCAAAATTTAGAACTTTAACTTTAAAGTTT 3505
Qy 1707 CTTCTGTAATTTGTAACCGGTGATTTATTAATCTCTAGATAAGCAGTACTAGAAACC 1766
Db 3506 CTTCTGTAATTTGTAACCGGTGATTTATTAATCTCTAGATAAGCAGTACTAGAAACC 3565
Qy 1767 AAACTCAGAAAAATTTTACCTGTTAGAAATTTCTTAAATTTTAAAGTGTGTTATCTTTT 1826
Db 3566 AAACTCAGAAAAATTTTACCTGTTAGAAATTTCTTAAATTTTAAAGTGTGTTATCTTTT 3625

```



```
QY 1827 CATTGGGTGATGTCAGGGTGATAAACAAGACATTCATCGAAGGCGATGCGATTGTGCCATT 1886
DB 3626 CATTGGGTGATGTCAGGGTGATAAACAAGACATTCATCGAAGGCGATGCGATTGTGCCATT 3685
QY 1887 GTGACAGTTTGTAAATAAACCACATACACACTTTTAAAGATTAAGTAAATCTAATCGA 1946
DB 3686 GTGACAGTTTGTAAATAAACCACATACACACTTTTAAAGATTAAGTAAATCTAATCGA 3745
QY 1947 AAGTCAGCTTGGAAATGGACATTTCCAAAGTATGTTTGGTGAAGTACACAGATATAAAATA 2006
DB 3746 AAGTCAGCTTGGAAATGGACATTTCCAAAGTATGTTTGGTGAAGTACACAGATATAAAATA 3805
QY 2007 GAAATCTGATGAGAGTTTTCAGTTTAAATACCAAGTCTTTAGGAGTCTTAACATTTGGC 2066
DB 3806 GAAATCTGATGAGAGTTTTCAGTTTAAATACCAAGTCTTTAGGAGTCTTAACATTTGGC 3865
QY 2067 CAGCATCTGTTTATCAATGACATAAATACGTTAAACCTTAAAGTTTAAATTTAATTT 2126
DB 3866 CAGCATCTGTTTATCAATGACATAAATACGTTAAACCTTAAAGTTTAAATTTAATTTAATTT 3925
QY 2127 AGGCAATTTATGCTGTGATTAATCTTACGGGAGAAAGAGATTTGATTGGAAAGCAGTT 2186
DB 3926 AGGCAATTTATGCTGTGATTAATCTTACGGGAGAAAGAGATTTGATTGGAAAGCAGTT 3985
QY 2187 TGGGAAGAAAGTGTCTGCAATTTCCAGAAATTTAAATGATTGGTTACAT 2236
DB 3986 TGGGAAGAAAGTGTCTGCAATTTCCAGAAATTTAAATGATTGGTTACAT 4035

RESULT 8
US-10-425-114A-26852
; Sequence 26852, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26852
; LENGTH: 4035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4654-056-G4_FLI
US-10-425-114A-26852

Query Match 84.6%; Score 1908.4; DB 54; Length 4035;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 327 ATATGAATCTGGACATTTGGGTGAAGGACCTTTGGCAAAAGTTGTAGAGTGCATTGA 386
DB 2126 AGATGAATCTGGACATTTGGGTGAAGGACCTTTGGCAAAAGTTGTAGAGTGCATTGA 2185
QY 387 TCATGGCATGATGCGATGATGAGTGAAGTAAATCGTAAATATGAGCCGTTACCG 446
DB 2186 TCATGGCATGATGCGATGATGAGTGAAGTAAATCGTAAATATGAGCCGTTACCG 2245
QY 447 TGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAATAGTACTGATCCCAATAG 506
DB 2246 TGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAATAGTACTGATCCCAATAG 2305
QY 507 TGTCTTCGATGTGCCAGATGCTAGAAATGGTTTGATCATCATGTTGATGTTGTTGTT 566
DB 2306 TGTCTTCGATGTGCCAGATGCTAGAAATGGTTTGATCATCATGTTGATGTTGTTGTTGTT 2365
```

```
QY 567 GTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACAGCTTTCTGCCATT 626
DB 2366 GTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACAGCTTTCTGCCATT 2425
QY 627 TCATAATGACACATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA 686
DB 2426 TCATAATGACACATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA 2485
QY 687 TCATAATAAATTAACCCCATACAGATCTGAAGCCTGAAATATTTTTTTGTGAAGTCTGA 746
DB 2486 TCATAATAAATTAACCCCATACAGATCTGAAGCCTGAAATATTTTTTTGTGAAGTCTGA 2545
QY 747 CTATGTAGTCAAAATATATCTTAAAAATGAACGTCATGAACGACACCTGAAAAACACAGA 806
DB 2546 CTATGTAGTCAAAATATATCTTAAAAATGAACGTCATGAACGACACCTGAAAAACACAGA 2605
QY 807 TATCAAGTTGTTGACCTTTGGAAGTGAACGTCATGAATGAATCAACAGTACTTTGGT 866
DB 2606 TATCAAGTTGTTGACCTTTGGAAGTGAACGTCATGAATGAATCAACAGTACTTTGGT 2665
QY 867 GTCTACCCGCGACCTACAGAGCTCCGAGGTCAATTTTGGCTTTAGTGTGCTCTCAGCCCTTG 926
DB 2666 GTCTACCCGCGACCTACAGAGCTCCGAGGTCAATTTTGGCTTTAGTGTGCTCTCAGCCCTTG 2725
QY 927 TGATGTTTGGAGCATAGGTTGCAATTTCTTATTAATATTTACCTTTGGTTTACAGTCTTTCA 986
DB 2726 TGATGTTTGGAGCATAGGTTGCAATTTCTTATTAATATTTACCTTTGGTTTACAGTCTTTCA 2785
QY 987 GACTCATGATAGTAAAGAGCACCTGGGCAATGATGAACGAATATTTAGGACCCATACCA 1046
DB 2786 GACTCATGATAGTAAAGAGCACCTGGGCAATGATGAACGAATATTTAGGACCCATACCA 2845
QY 1047 ACACATGATTCAGAAAAACAAGAACGCAAGTATTTTCCACCAACACAGCTAGATTGGA 1106
DB 2846 ACACATGATTCAGAAAAACAAGAACGCAAGTATTTTCCACCAACACAGCTAGATTGGA 2905
QY 1107 TGAACACAGATTTCTGCTGGTAGATATGTTTAGGAGACGCTGCAAAACCGTTTGAAGGAATTTAT 1166
DB 2906 TGAACACAGATTTCTGCTGGTAGATATGTTTAGGAGACGCTGCAAAACCGTTTGAAGGAATTTAT 2965
QY 1167 GCTTTGTCATGATGAAGAACATGAGAACTGTTTGAACCTGGTTGGAAGAAATTTAGAATA 1226
DB 2966 GCTTTGTCATGATGAAGAACATGAGAACTGTTTGAACCTGGTTGGAAGAAATTTAGAATA 3025
QY 1227 TGATCCCACTCAAGAAATTAACCTTTGGATGAAGATGCGAGCATCTTTTCTTTGACTTAT 1286
DB 3026 TGATCCCACTCAAGAAATTAACCTTTGGATGAAGATGCGAGCATCTTTTCTTTGACTTAT 3085
QY 1287 AAAAAAGAAATGAAATGGGAATCAGTGTCTTACTATATATCTTCTAGAGAGAGATTACT 1346
DB 3086 AAAAAAGAAATGAAATGGGAATCAGTGTCTTACTATATATCTTCTAGAGAGAGATTACT 3145
QY 1347 TAAGACTGTGTGCTCAACTAAACATTTCTAATAATTTTTTGTAAACATTAATAATTTTGTGA 1406
DB 3146 TAAGACTGTGTGCTCAACTAAACATTTCTAATAATTTTTTGTAAACATTAATAATTTTGTGA 3205
QY 1407 CAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATAAATTAACCTTTTGAAGCAAGT 1466
DB 3206 CAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATAAATTAACCTTTTGAAGCAAGT 3265
QY 1467 ATGCTCTTGATAATGCAATTAAGAAAAATTAATAATTTTTTCTTTTGAATTTACCAATTT 1526
DB 3266 ATGCTCTTGATAATGCAATTAAGAAAAATTAATAATTTTTTCTTTTGAATTTACCAATTT 3325
QY 1527 TTAATACCTTTTGAATAATCTTTTGTGTCAGTGAATAAATGATGATGATGCTGCTTTG 1586
DB 3326 TTAATACCTTTTGAATAATCTTTTGTGTCAGTGAATAAATGATGATGATGCTGCTTTG 3385
QY 1587 TACATGGAGGTCACTCTGAGTGTATTTTTTTTGGTGAAGAAAGAAATCTTCACTACTTTA 1646
DB 3386 TACATGGAGGTCACTCTGAGTGTATTTTTTTTGGTGAAGAAAGAAATCTTCACTACTTTA 3445
```

QY 1647 TATTCCTTAAGGAATATCTTTTATATATCTCAAAATTTAGAACTTAAACTTTAAAAGTTT 1706  
 Db 3446 TATTCCTTAAGGAATATCTTTTATATATCTCAAAATTTAGAACTTAAACTTTAAAAGTTT 3505  
 QY 1707 CTTCTGTAATTTGTAACCGGTGATTTATTAATCTCTAGATAAGCAGGTACTAGAAACC 1766  
 Db 3506 CTTCTGTAATTTGTAACCGGTGATTTATTAATCTCTAGATAAGCAGGTACTAGAAACC 3565  
 QY 1767 AAACTCAGAAATCTGTTACTGTTAGAAATCTATTAATAATTTTAAAGTGTGTTATCTTTT 1826  
 Db 3566 AAACTCAGAAATCTGTTACTGTTAGAAATCTATTAATAATTTTAAAGTGTGTTATCTTTT 3625  
 QY 1827 CATTTGGGTGATCTCAGGTGTATTAACAGACATTTATGAAAGGCATGCAAGTTTGTCCATT 1886  
 Db 3626 CATTTGGGTGATCTCAGGTGTATTAACAGACATTTATGAAAGGCATGCAAGTTTGTCCATT 3685  
 QY 1887 GTGACAGTTTGTATTAATAAACCAATACATACATCTTTTATTAAGATTAATAATCTAACTGGA 1946  
 Db 3686 GTGACAGTTTGTATTAATAAACCAATACATACATCTTTTATTAAGATTAATAATCTAACTGGA 3745  
 QY 1947 AAGTCAGCTTGGAAATCGACATTTCCAAAGTATGTTTGGTGAATGTCACAGATATAAAATA 2006  
 Db 3746 AAGTCAGCTTGGAAATCGACATTTCCAAAGTATGTTTGGTGAATGTCACAGATATAAAATA 3805  
 QY 2007 GAAATCTGATGAGAGGTTTCAGTTTAAATACCAAGTCTTTAGAGTCTTTAAACATTCGC 2066  
 Db 3806 GAAATCTGATGAGAGGTTTCAGTTTAAATACCAAGTCTTTAGAGTCTTTAAACATTCGC 3865  
 QY 2067 CAGCATCTGTTTATCAATGACATTAATACGTAACCTTAAAGATTAAGTTTATTAATT 2126  
 Db 3866 CAGCATCTGTTTATCAATGACATTAATACGTAACCTTAAAGATTAAGTTTATTAATT 3925  
 QY 2127 AGGCAATTTATGTCGTGATTAATCTTACGGGAGAAAGAGGATTTGATTGGAAGCAGTT 2186  
 Db 3926 AGGCAATTTATGTCGTGATTAATCTTACGGGAGAAAGAGGATTTGATTGGAAGCAGTT 3985  
 QY 2187 TGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATGTTACAT 2236  
 Db 3986 TGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATGTTACAT 4035

RESULT 9

PCT-US01-14827-2154  
 ; Sequence 2154, Application PC/TUS0114827  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hysq, Inc  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: 21272-104  
 ; CURRENT APPLICATION NUMBER: PCT/US01/14827  
 ; CURRENT FILING DATE: 2001-05-16  
 ; PRIOR APPLICATION NUMBER: 09/577,408  
 ; PRIOR FILING DATE: 2000-05-18  
 ; NUMBER OF SEQ ID NOS: 16102  
 ; SOFTWARE: Custom  
 ; SEQ ID NO 2154  
 ; LENGTH: 2494  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SIMILAR  
 ; LOCATION: (81)..(203)  
 ; OTHER INFORMATION: 92% homologous to Homo sapiens Novel protein kinase, SEQ ID  
 ; OTHER INFORMATION: NO: 175, accession number AAB5648, Smith-Waterman Score=213.  
 PCT-US01-14827-2154

Query Match 80.9%; Score 1825; DB 1; Length 2494;  
 Best Local Similarity 92.4%; Pred. No. 0;  
 Matches 2093; Conservative 0; Mismatches 60; Indels 112; Gaps 12;

QY 14 GTCATTATTAGAACGAGTCTTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 73  
 Db 309 GTCATTATTAGAACGAGTCTTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 368

QY 74 ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATATATCACAGACATTTG 133  
 Db 369 ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATATATCACAGACATTTG 428  
 QY 134 AAACCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGAGCAGGAGAGCAGTCTCTA 193  
 Db 429 AAACCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGAGCAGGAGAGCAGTCTCTA 488  
 QY 194 AAAGGAAGCGCAATAGACACATGTTCAAGTCAATCAGTCACTGTTTGAAGAGCCACCGAAGGA 253  
 Db 489 AAAGGAAGCGCAATAGACACATGTTCAAGTCAATCAGTCACTGTTTGAAGAGCCACCGAAGGA 536  
 QY 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAAAAGTGAGAGCG 313  
 Db 537 ----- 536  
 QY 314 TTCTAAGAGCAAGATATGNAATCGTGGACACTTTGGGTGAAGGAGCCCTTTGCGCAAGTTG 373  
 Db 537 -----AATCCGTGGACACTTTGGGTGAAGGAGCCCTTTGCGCAAGTTG 578  
 QY 374 TAGAGTGCATTGATCATGGCATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATG 433  
 Db 579 TAGAGTGCATTGATCATGGCATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATG 638  
 QY 434 TAGGCGGTTACCGTGAAGCAGCTCGTTCAAGAAATCCAAAGTATTTAGAGCAGTAAATAGTA 493  
 Db 639 TAGGCGGTTACCGTGAAGCAGCTCGTTCAAGAAATCCAAAGTATTTAGAGCAGTAAATAGTA 698  
 QY 494 CTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAAATGTTTGTGATCATCATGTC 553  
 Db 699 CTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAAATGTTTGTGATCATCATGTC 758  
 QY 554 ATGTTTGTATTTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 613  
 Db 759 ATGTTTGTATTTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 818  
 QY 614 GCTTTCTGCCATTTCAAATTTGACCAATGAGCAGATGCGCGTATCAGATCTGCCAGTCAA 673  
 Db 819 GCTTTCTGCCATTTCAAATTTGACCAATGAGCAGATGCGCGTATCAGATCTGCCAGTCAA 878  
 QY 674 TAAATTTTTACATCATATAAATTTAACCATACAGATCTGAAGCTCAAAATATTTTGT 733  
 Db 879 TAAATTTTTACATCATATAAATTTAACCATACAGATCTGAAGCTCAAAATATTTTGT 938  
 QY 734 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAATGAAACGCTGATGAACACAC 998  
 Db 999 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAATGAAACGCTGATGAACACAC 998  
 QY 794 TGAATAACACAGATATCAAAGTGTGACTTTGGAGTGCACCTGATGATGATGAACATC 853  
 Db 999 TGAATAACACAGATATCAAAGTGTGACTTTGGAGTGCACCTGATGATGATGAACATC 1058  
 QY 854 ACAGTACTTTGTTGTTACCCGCACTACAGAGTCTCCGAGTCAATTTGGCTTTAGGTT 913  
 Db 1059 ACAGTACTTTGTTGTTACCCGCACTACAGAGTCTCCGAGTCAATTTGGCTTTAGGTT 1118  
 QY 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGCTTATTGAAATATTAATCCCTGGG 971  
 Db 1119 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGCTTATTGAAATATTAATCCCTGGG 1178  
 QY 972 TTTCAAGTCTTTGAGAC-TCATGATGATGAAGAGCCTCGCAATGATGAACGAATAT 1030  
 Db 1179 GTTCAAGTCTTTGAGAC-TCATGATGATGAAGAGCCTCGCAATGATGAACGAATAT 1238  
 QY 1031 TAGGACCCCAT-CCACAAACACATGATT-CAGAAAAACAAGAAACCGAAGTATTTTCCAC 1086  
 Db 1239 TTAGGACCCCATACCCCAACACATGATTCCAGAAAAACAAGAAACCGAAGTATTTTCCAC 1298  
 QY 1087 C---ATAACAGCTAGATTGGATGAACACAGTTTC--TGCTGTGTAGATATGTTAGGAGAC 1141  
 Db 1299 CCATAACCCAGCCTAGATTGGATGAACACAGTTTCCTGCTGTGTAGATATGTTAGGAGAC 1358  
 QY 1142 GCTGCAACCCGTTGAAGAAATTTATGCTTTGTCATGATGAAGAACATGAGAAATCTGTTTG 1201



Qy 614 GCTTTCTGCCATTTCAAAATTCACACATCAGCGAGATGGCGTATCAGATCTGCCAGTCAA 673  
 Db 819 GCTTTCTGCCATTTCAAAATTCACACATCAGCGAGATGGCGTATCAGATCTGCCAGTCAA 878  
 Qy 674 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 733  
 Db 879 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 938  
 Qy 734 TTGTGAAGTCTGACTATGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 793  
 Db 939 TTGTGAAGTCTGACTATGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 998  
 Qy 794 TGAATAACACAGATATCAAAATTTGTGACTTTTGAAGTGCACATGATGATCAATCAATC 853  
 Db 999 TGAATAACACAGATATCAAAATTTGTGACTTTTGAAGTGCACATGATGATCAATCAATC 1058  
 Qy 854 ACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 913  
 Db 1059 ACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 1118  
 Qy 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGATCTTATTTGAATATTA--CCTTGG 971  
 Db 1119 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGATCTTATTTGAATATTAATACCCCTGG 1178  
 Qy 972 TTTTCACAGTCTTTCAGAC--TCATGATAGTAAGAGCCTCGCAATGATGGAACGAATAT 1030  
 Db 1179 GTTTCACAGTCTTTCAGACCTTCAATGATGAAGAGCCTCGCAATGATGGAACGAATAT 1238  
 Qy 1031 TAGGACCCCATTA--CCACAAACACATGATT--CAGAAAAAAGAAACGCAAGTATTTTTCAC 1086  
 Db 1239 TTAGGACCCCATTA--CCACAAACACATGATTCCAGAAAAAAGAAACGCAAGTATTTTTCAC 1298  
 Qy 1087 C---ATAACAGCTAGATTGGATGAACACAGTTC--TGCTGGTAGATATGTTTAGGAGAC 1141  
 Db 1299 CCATAACCCAGCCTAGATTGGATGAACACAGTTCCTGCTGGTAGATATGTTTAGGAGAC 1358  
 Qy 1142 GCTGCAAAACCGTTGAAGAAATTTATGCTTTCATGATGAAGAAACATGAGAAATCTGTTG 1201  
 Db 1359 GCTGCAAAACCGTTGAAGAAATTTATGCTTTCATGATGAAGAAACATGAGAAATCTGTTG 1418  
 Qy 1202 ACCTGGTTGCAAGAAATTTAGATATGATCAACTCAAAAGAAATTAACCTTGGATGAAGAT 1261  
 Db 1419 ACCTGGTTGCAAGAAATTTAGATATGATCAACTCAAAAGAAATTAACCTTGGATGAAGAT 1478  
 Qy 1262 TGCAGCATCTTTCTTTCAGCTTTTAAAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1321  
 Db 1479 TGCAGCATCTTTCTTTCAGCTTTTAAAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1538  
 Qy 1322 ATATACCTTCTAGAGAGATTAACCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1381  
 Db 1539 ATATACCTTCTAGAGAGATTAACCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1598  
 Qy 1382 TTTGTAACATTAATTTTGTACAGTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1441  
 Db 1599 TTTGTAACATTAATTTTGTACAGTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1658  
 Qy 1442 AGCATAATTAATTTTGAAGCAAGTATGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1501  
 Db 1659 AGCATAATTAATTTTGAAGCAAGTATGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1718  
 Qy 1502 ATTTTCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1561  
 Db 1719 ATTTTCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1778  
 Qy 1562 TAAATGTGATGATCTTGCCTTTTGTACATGAGAGTCACTCTGAGAGTGAATTTTGTGA 1621  
 Db 1779 TAAATGTGATGATCTTGCCTTTTGTACATGAGAGTCACTCTGAGAGTGAATTTTGTGA 1838  
 Qy 1622 GTAAAGGAATCTTGACTTCTTATATCTTAAAGGAATATCTTAAAGGAATCTTAAAGGAATCTTAAAGGAAT 1681  
 Db 1839 GTAAAGGAATCTTGACTTCTTATATCTTAAAGGAATATCTTAAAGGAATCTTAAAGGAATCTTAAAGGAAT 1898

Qy 1682 TTAGAACTTAACCTTTAAAGATTTTCTCTGTAATTTGTTGAACGGGTGATTAATTAATTAAC 1741  
 Db 1899 TTAGAACTTAACCTTTAAAGATTTTCTCTGTAATTTGTTGAACGGGTGATTAATTAATTAAC 1958  
 Qy 1742 TCTAGATAGCAGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTTATT 1801  
 Db 1959 TCTAGATAGCAGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTTATT 2018  
 Qy 1802 AAATTTTAAAGTGTGATTTCTTTTCAATTTGAGTGTGAGGTGATTAAC--AGACATT 1859  
 Db 2019 AAATTTTAAAGTGTGATTTCTTTTCAATTTGAGTGTGAGGTGATTAACCCAGACATT 2078  
 Qy 1860 CATGGAAGG---CATGCACTTTGTCCATTTGACAGTGTGTTTAAATAAACCAAC--ATAC 1915  
 Db 2079 CATGGAAGGCGATGCCAGTTTGTCCATTTGACAGTGTGTTTAAATAAACCAACGATAC 2138  
 Qy 1916 ACATTTTATTAAGATTAATTAATC--TAACTGGAAGTCACTTGGAAATGGAATTTTCC 1973  
 Db 2139 ACATTTTATTAAGATTAATTAATCCTGTAAGGTGAGTGTGGAATTTGGAATTTTCC 2198  
 Qy 1974 AAGTATGTTTGGTGAAGTCAAGATATAAAATAGAAATTTCTGATGAGAGTGTTCAGTTTT 2033  
 Db 2199 AAGTATGTTTGGTGAAGTCAAGATATAAAATAGAAATTTCTGATGAGAGTGTTCAGTTTT 2258  
 Qy 2034 TAATACCAAGTCTTCTAGGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAA 2093  
 Db 2259 TAATACCAAGTCTTCTAGGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAA 2318  
 Qy 2094 TAGTAACCTTAAGAAATTT--AAGTTTATTAATTAAGCAATTTATGCTGTGATTAATTC 2151  
 Db 2319 TAGTAACCTTAAGAAATTTAAGGTTTATTAATTAAGCAATTTATGCTGTGATTAATTC 2378  
 Qy 2152 TTACGGGAGAAAGAGGATTTGATTGGAAGACAGTGTGGGAAGAAAGTGTCTGTGAAATTT 2211  
 Db 2379 TTACGGGAGAAAGAGGATTTGATTGGAAGACAGTGTGGGAAGAAAGTGTCTGTGAAATTT 2438  
 Qy 2212 CCAGAAATTTAATTTGATTTGTTTACATAAACTTTTGTGATTTCCAGAAA 2256  
 Db 2439 CCAGAAATTTAATTTGATTTGTTTACATAAACTTTTGTGATTTCCAGAAA 2483

RESULT 11  
 US-10-170-235-20526  
 ; Sequence 20526, Application US/10170235  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig  
 ; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN  
 ; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF  
 ; FILE REFERENCE: C1001380  
 ; CURRENT APPLICATION NUMBER: US/10/170,235  
 ; CURRENT FILING DATE: 2003-03-17  
 ; NUMBER OF SEQ ID NOS: 42514  
 ; SEQ ID NO 20526  
 ; LENGTH: 1795  
 ; TYPE: DNA  
 ; ORGANISM: HUMAN  
 ; US-10-170-235-20526

Query Match 72.58; Score 1635; DB 49; Length 1795;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GTCAATTTTAGAAGCAAGGTCTTTGAATGAGCGAGATATTCGGGACCGGAGATACGTTG 73  
 Db 161 GTCAATTTTAGAAGCAAGGTCTTTGAATGAGCGAGATATTCGGGACCGGAGATACGTTG 220  
 Qy 74 ACGAATACAGAATGACTCTGTGAAGATATGTTCTTAGACATTTACAGAGACATTG 133  
 Db 221 ACGAATACAGAATGACTCTGTGAAGATATGTTCTTAGACATTTACAGAGACATTG 280  
 Qy 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGGAGAGCAAGTCTCTA 193  
 Db 281 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGGAGAGCAAGTCTCTA 340

```
QY 194 AAAGGAGCGCAATAGACACTGTTCAAGTCATCACTGCTCGAAGAGCCACCGAAGGA 253
Db |||
QY 341 AAAGGAGCGCAATAGACACTGTTCAAGTCATCACTGCTCGAAGAGCCACCGAAGGA 400
Db |||
QY 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCCTGATCTCTCAAGTGAGAGG 313
Db |||
QY 401 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCCTGATCTCTCAAGTGAGAGG 460
Db |||
QY 314 TTCTAAGAGCAAGATATGAATCGTGGACACATTTGGGTGAAGGAGCCTTTGGCAAGTTG 373
Db |||
QY 461 TTCTAAGAGCAAGATATGAATCGTGGACACATTTGGGTGAAGGAGCCTTTGGCAAGTTG 520
Db |||
QY 374 TAGAGTGCAATTGATCATGGCATGGATGCGATGTCATGTAGCAGTGAAGATCGTAAAGATG 433
Db |||
QY 521 TAGAGTGCAATTGATCATGGCATGGATGCGATGTCATGTAGCAGTGAAGATCGTAAAGATG 580
Db |||
QY 434 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAGATATTAAGTAGTA 493
Db |||
QY 581 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAGATATTAAGTAGTA 640
Db |||
QY 494 CTGATCCCAATAGTGTCTTCCGATGCTCCAGATGCTAGAAATGTTGATCATCATGTC 553
Db |||
QY 641 CTGATCCCAATAGTGTCTTCCGATGCTCCAGATGCTAGAAATGTTGATCATCATGTC 700
Db |||
QY 554 ATGTTTGTATTGTTGAACTACTGGGACCTTAGTACTTACGATTTCAATTAAGAAACA 613
Db |||
QY 701 ATGTTTGTATTGTTGAACTACTGGGACCTTAGTACTTACGATTTCAATTAAGAAACA 760
Db |||
QY 614 GCTTCTGCGATTTCAAATGACCATCAGCAGATGCGGATCAGATCTGCCAGTCAA 673
Db |||
QY 761 GCTTCTGCGATTTCAAATGACCATCAGCAGATGCGGATCAGATCTGCCAGTCAA 820
Db |||
QY 674 TAAATTTTTCATCATATAATTAATTAACCATACAGATCTGAAGCTGAAATATTTTGT 733
Db |||
QY 821 TAAATTTTTCATCATATAATTAATTAACCATACAGATCTGAAGCTGAAATATTTTGT 880
Db |||
QY 734 TTGGAAGTCTGACTATGATGATCAATATAATTTCTAAATGAACGCTGATGAACGACAC 793
Db |||
QY 881 TTGGAAGTCTGACTATGATGATCAATATAATTTCTAAATGAACGCTGATGAACGACAC 940
Db |||
QY 794 TGAAGAACACAGATATCAAGTTGTTGACTTGGAGTCCGAGTCAATGATGATGAACATC 853
Db |||
QY 941 TGAAGAACACAGATATCAAGTTGTTGACTTGGAGTCCGAGTCAATGATGATGAACATC 1000
Db |||
QY 854 ACAGTACTTTGTTGTTGTTGAGCATAGTTGCTTCAATTTTCAATTAATTAATTAATTAAT 913
Db |||
QY 1001 ACAGTACTTTGTTGTTGTTGAGCATAGTTGCTTCAATTTTCAATTAATTAATTAATTAAT 1060
Db |||
QY 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGTTGCTTCAATTTTCAATTAATTAATTAATTAAT 973
Db |||
QY 1061 GGTCTCAGCCTTTGATGTTTGGAGCATAGTTGCTTCAATTTTCAATTAATTAATTAATTAAT 1120
Db |||
QY 974 TCAGAGTCTTTGATGTTTGGAGCATAGTTGCTTCAATTTTCAATTAATTAATTAATTAAT 1033
Db |||
QY 1121 TCAGAGTCTTTGATGTTTGGAGCATAGTTGCTTCAATTTTCAATTAATTAATTAATTAAT 1180
Db |||
QY 1034 GACCCATACCAACACATGATTTCAAGAAACAGGAAACGCAAGTATTTTCAACATACAC 1093
Db |||
QY 1181 GACCCATACCAACACATGATTTCAAGAAACAGGAAACGCAAGTATTTTCAACATACAC 1240
Db |||
QY 1094 AGCTAGATTTGGATGAACACATGTTCTGCTGGTAGATATGTTAGGAGAGCTGCAACCGT 1153
Db |||
QY 1241 AGCTAGATTTGGATGAACACATGTTCTGCTGGTAGATATGTTAGGAGAGCTGCAACCGT 1300
Db |||
QY 1154 TGAAGGAATTTATGCTTTGATGATGAAGAAACATGAGAAATCTGTTGACCTGTTGCAAA 1213
Db |||
QY 1301 TGAAGGAATTTATGCTTTGATGATGAAGAAACATGAGAAATCTGTTGACCTGTTGCAAA 1360
Db |||
QY 1214 GAATGTTAGATATGATCAACTCAAGAAATTTACCTTGGATGAAGCAATTCGAGCATCCTT 1273
Db |||
QY 1361 GAATGTTAGATATGATCAACTCAAGAAATTTACCTTGGATGAAGCAATTCGAGCATCCTT 1420
Db |||
```

```
QY 1274 TCTTTGACTTATTAATAAAGAAATGAAATGGAATCAGTGGTCTTATCTATATATCTTCT 1333
Db |||
QY 1421 TCTTTGACTTATTAATAAAGAAATGAAATGGAATCAGTGGTCTTATCTATATATCTTCT 1480
Db |||
QY 1334 AGAAGAGATTAATTAAGAGTCTGTCAGTCAATCAATCAATCTTAATATTTTGTAAACAT 1393
Db |||
QY 1481 AGAAGAGATTAATTAAGAGTCTGTCAGTCAATCAATCAATCTTAATATTTTGTAAACAT 1540
Db |||
QY 1394 AAATTAATTTTGTACAGTTAAAGTGAATATTTGTATGTTTGTATCAATAGCATATAAT 1453
Db |||
QY 1541 AAATTAATTTTGTACAGTTAAAGTGAATATTTGTATGTTTGTATCAATAGCATATAAT 1600
Db |||
QY 1454 TTGTTAAGCAAGTATGTCGTTGATTAATCAATAGCAATTAATAATTTTCTTTT 1513
Db |||
QY 1601 TTGTTAAGCAAGTATGTCGTTGATTAATCAATAGCAATTAATAATTTTCTTTT 1660
Db |||
QY 1514 GAAATTAACATTTTAAATACCTTTGAATATCTTTGTCAGTGAATATGATGATG 1573
Db |||
QY 1661 GAAATTAACATTTTAAATACCTTTGAATATCTTTGTCAGTGAATATGATGATG 1720
Db |||
QY 1574 ATCTTGCTTTTGTACATGAGGTCACCTCTGAACTGATTTTGTAGTAAAGAAAT 1633
Db |||
QY 1721 ATCTTGCTTTTGTACATGAGGTCACCTCTGAACTGATTTTGTAGTAAAGAAAT 1780
Db |||
QY 1634 CTGACTACTTTATA 1648
Db |||
QY 1781 CTGACTACTTTATA 1795
Db |||
```

## RESULT 12

US-10-932-349-580

; Sequence 580, Application US/10932349

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: ENCODING HUMAN KINASE PROTEINS, METHODS OF DETECTION AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL001483

; CURRENT APPLICATION NUMBER: US/10/932,349

; CURRENT FILING DATE: 2004-09-02

; NUMBER OF SEQ ID NOS: 123188

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 580

; LENGTH: 1795

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-932-349-580

Query Match 72.5%; Score 1634.6; DB 65; Length 1795;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1634; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 14 GTCATTATTTAGACGAAGTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 73
Db |||
QY 161 GTCATTATTTAGACGAAGTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 220
Db |||
QY 74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCAGAGACATTTG 133
Db |||
QY 221 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCAGAGACATTTG 280
Db |||
QY 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCTTA 193
Db |||
QY 281 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCTTA 340
Db |||
QY 194 AAAGGAGCGCAATAGACACTGTTCAAGTCATCACTGATGATGATGATGATGATGATGATG 253
Db |||
QY 341 AAAGGAGCGCAATAGACACTGTTCAAGTCATCACTGATGATGATGATGATGATGATGATG 400
Db |||
QY 254 AAAGATCCAGGAGTATGAGGATGATGAGGAGGTCACCTGATCTGTCATGATGATGATGATG 313
Db |||
QY 401 AAAGATCCAGGAGTATGAGGATGATGAGGAGGTCACCTGATCTGTCATGATGATGATGATG 460
Db |||
QY 314 TTCTAAGAGCAAGATATGAATCGTGGACACTTTTGGGTGAAGGAGCCTTTTGGCAAGTTG 373
Db |||
```



QY 494 CTGATCCCAATAGTGTCTTCGATGTGCCAGATGCTAGAAATGGTTTGCATCATCATGTC 553  
Db 641 CTGATCCCAATAGTGTCTTCGATGTGCCAGATGCTAGAAATGGTTTGCATCATCATGTC 700  
QY 554 ATGTTTGTATTTGTTGAACTTACCTGGGACCTAGTACTTACGATTTTCAATTAAGAAACA 613  
Db 701 ATGTTTGTATTTGTTGAACTTACCTGGGACCTAGTACTTACGATTTTCAATTAAGAAACA 760  
QY 614 GCTTTCTGCCATTTCAAATTTGACACATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA 673  
Db 761 GCTTTCTGCCATTTCAAATTTGACACATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA 820  
QY 674 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCTGAAAATATTTTGT 733  
Db 821 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCTGAAAATATTTTGT 880  
QY 734 TTGTGAAGTCTGACTATGTAGTCAATATATTTCAAATGAAAGCTGATGAACGCACAC 793  
Db 881 TTGTGAAGTCTGACTATGTAGTCAATATATTTCAAATGAAAGCTGATGAACGCACAC 940  
QY 794 TGAATAACACAGATATCAAAGTTGTGACTTTGGAAGTGCACAGTATGATGAACATC 853  
Db 941 TGAATAACACAGATATCAAAGTTGTGACTTTGGAAGTGCACAGTATGATGAACATC 1000  
QY 854 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGTCTATTTTGGCTTTAGGTT 913  
Db 1001 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGTCTATTTTGGCTTTAGGTT 1060  
QY 914 GGTCTCAGCTTGTGATGTTGGAGCATAGTTGCAATCTTATTTGAATATTAACCTTGGTT 973  
Db 1061 GGTCTCAGCTTGTGATGTTGGAGCTAGTTGCAATCTTATTTGAATATTAACCTTGGTT 1120  
QY 974 TCACAGTCTTCAGACTCATGATGATGAAGACCTCGCATGATGAACGAATATTAG 1033  
Db 1121 TCACAGTCTTCAGACTCATGATGATGAAGACCTCGCATGATGAACGAATATTAG 1180  
QY 1034 GACCCATACCAACACACATGATTTCAGAAAAACGAAAGTATTTTCAACATAACC 1093  
Db 1181 GACCCATACCAACACACATGATTTCAGAAAAACGAAAGTATTTTCAACATAACC 1240  
QY 1094 AGCTAGATGGGATGAACACAGTTCTGCTGTGTAGATATGTTAGAGACGCTGCACAAACCGT 1153  
Db 1241 AGCTAGATGGGATGAACACAGTTCTGCTGTGTAGATATGTTAGAGACGCTGCACAAACCGT 1300  
QY 1154 TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAAACTGTTTGCACCTGGTTGCA 1213  
Db 1301 TGAAGGAATTTATGCTTTTGTATGATGAAGAACATGAGAAACTGTTTGCACCTGGTTGCA 1360  
QY 1214 GAAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGACAGATCCTT 1273  
Db 1361 GAAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGACAGATCCTT 1420  
QY 1274 TCTTTGACTTATTAAGAAAGTAAATGAAATGGGAATCAGTGTCTTATATATCTCTCT 1333  
Db 1421 TCTTTGACTTATTAAGAAAGTAAATGAAATGGGAATCAGTGTCTTATATATCTCTCT 1480  
QY 1334 AGAGAGATTAATTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTT 1393  
Db 1481 AGAGAGATTAATTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTT 1540  
QY 1394 AAAATTTTGTACAGTTTAAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC 1453  
Db 1541 AAAATTTTGTACAGTTTAAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC 1600  
QY 1454 TTGTTAAGCAAGTATGGTCTTGATATGCAATAGAAAAATTAATAATTTTCTTTT 1513  
Db 1601 TTGTTAAGCAAGTATGGTCTTGATATGCAATAGAAAAATTAATAATTTTCTTTT 1660  
QY 1514 GAAATTTACATTTTAAATATCTTTGAAATATCTTTTGTGTCAGTGAATAATGATG 1573  
Db 1661 GAAATTTACATTTTAAATATCTTTGAAATATCTTTTGTGTCAGTGAATAATGATG 1720

QY 1574 ATCTTGCCCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTAGTAAAGAAAT 1633  
Db 1721 ATCTTGCCCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTAGTAAAGAAAT 1780  
QY 1634 CTTGACTACTTTTATA 1648  
Db 1781 CTTGACTACTTTTATA 1795  
RESULT 14  
US-10-756-149-49  
; Sequence 49, Application US/10756149  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND  
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER  
; FILE REFERENCE: file  
; CURRENT APPLICATION NUMBER: US/10/756.149  
; CURRENT FILING DATE: 2004-01-12  
; NUMBER OF SEQ ID NOS: 5818  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 49  
; LENGTH: 1814  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1562)..(1562)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-756-149-49

Query Match 72.3%; Score 1632; DB 62; Length 1814;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1638; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 14 GTCATTTTAAAGCAAGTCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 73  
Db 161 GTCATTTTAAAGCAAGTCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 220  
QY 74 ACGAATACAGGAATGACTCTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTG 133  
Db 221 ACGAATACAGGAATGACTCTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTG 280  
QY 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGACGAGGAGAGACGCTCTTA 193  
Db 281 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGACGAGGAGAGACGCTCTTA 340  
QY 194 AAAGGAAGCGCAATAGACATCTGTTCAAAGTCACTAGTCACTGTTCCGAAGAGCCACCGAAGGA 253  
Db 341 AAAGGAAGCGCAATAGACATCTGTTCAAAGTCACTAGTCACTGTTCCGAAGAGCCACCGAAGGA 400  
QY 254 AAAGTCCAGAGATATAGAGATGATGAGGAGGTCACTGATCTGTCAAAGTGGAGAGC 313  
Db 401 AAAGTCCAGAGATATAGAGATGATGAGGAGGTCACTGATCTGTCAAAGTGGAGAGC 460  
QY 314 TTCTTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTTGGCAAAGTTG 373  
Db 461 TTCTTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTTGGCAAAGTTG 520  
QY 374 TAGAGTGCATTTGATCATGGCATGGCATGCGCATGTCATGTAGCAGTGAATAATCGTAAATAATG 433  
Db 521 TAGAGTGCATTTGATCATGGCATGGCATGCGCATGTCATGTAGCAGTGAATAATCGTAAATAATG 580  
QY 434 TAGGCGGTACCGTGAAGCAGCTCGTTTCAAGAAATCCAAAGTATTAGAGCAGCTTAATAATGTA 493  
Db 581 TAGGCGGTACCGTGAAGCAGCTCGTTTCAAGAAATCCAAAGTATTAGAGCAGCTTAATAATGTA 640  
QY 494 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGTTTTCATCATCATGTC 553  
Db 641 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGTTTTCATCATCATGTC 700  
QY 554 ATGTTTGTATTTGTTTGAACACTACTGGGACTTAGTACTTACGATTTTTCATTAAGAAACA 613



Db 701 ATGTTTGTATGTTGTTGAACTACTGGGACTAGTACTTACGATTTCAATTAAGAAACA 760  
QY 614 GCTTTCTGCCATTTCAAAATGACCAATCAGCGAGATGGCGTATCAGATCTGCCAGTCAA 673  
Db 761 GCTTTCTGCCATTTCAAAATGACCAATCAGCGAGATGGCGTATCAGATCTGCCAGTCAA 820  
QY 674 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 733  
Db 821 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 880  
QY 734 TTGTGAAGCTGACATGATGATCAATATAATTTTAAATGAAACGTCATGAAGCAGCAC 793  
Db 881 TTGTGAAGCTGACATGATGATCAATATAATTTTAAATGAAACGTCATGAAGCAGCAC 940  
QY 794 TGAATAACACAGATATCAAGTGTGTTGACTTTGGAGTGCACGTCATGATGATGACATC 853  
Db 941 TGAATAACACAGATATCAAGTGTGTTGACTTTGGAGTGCACGTCATGATGATGACATC 1000  
QY 854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 913  
Db 1001 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 1060  
QY 914 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGCAATTTTATTAATTAATTAATTTGTT 973  
Db 1061 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGCAATTTTATTAATTAATTAATTTGTT 1120  
QY 974 TCACAGTCTTTCAGACTCATGATAGTAAGGACCTGGCAATGATGAAAGATATTAAG 1033  
Db 1121 TCACAGTCTTTCAGACTCATGATAGTAAGGACCTGGCAATGATGAAAGATATTAAG 1180  
QY 1034 GACCCATACCAACACACATGATTTAGAAAAAGCAAGTATTTTACACATAACC 1093  
Db 1181 GACCCATACCAACACACATGATTTAGAAAAAGCAAGTATTTTACACATAACC 1240  
QY 1094 AGCTAGATTTGGATGAAACACAGTCTGCTGTGTAGATATTTAGGAGACGTCGAAACCGT 1153  
Db 1241 AGCTAGATTTGGATGAAACACAGTCTGCTGTGTAGATATTTAGGAGACGTCGAAACCGT 1300  
QY 1154 TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAACTGTTTCACTCTGTTGCA 1213  
Db 1301 TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAACTGTTTCACTCTGTTGCA 1360  
QY 1214 GAATCTTGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCTT 1273  
Db 1361 GAATCTTGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCTT 1420  
QY 1274 TCTTTGACTTATTAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTCT 1333  
Db 1421 TCTTTGACTTATTAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTCT 1480  
QY 1334 AGAGAGATTTACTTAAGACTGTGTCAGTCACTAAACATTTCTAATATTTTGTAAACATT 1393  
Db 1481 AGAGAGATTTACTTAAGACTGTGTCAGTCACTAAACATTTCTAATATTTTGTAAACATT 1540  
QY 1394 AAATTAATTTGTACAGTTAAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC 1453  
Db 1541 AAATTAATTTGTACAGTTAAAGTGAATATTTGTATGTTTGTATCAATAGCATATTAAC 1600  
QY 1454 TTGTTAAGCAATGATGTTCTGATATGCAATAGAAAAATTAATTAATTTTCTTTTCT 1513  
Db 1601 TTGTTAAGCAATGATGTTCTGATATGCAATAGAAAAATTAATTAATTTTCTTTTCT 1660  
QY 1514 GAAATTTACATTTTAAATACCTTTGAAATACCTTTGTGTCAGTGAATATGATG 1573  
Db 1661 GAAATTTACATTTTAAATACCTTTGAAATACCTTTGTGTCAGTGAATATGATG 1720  
QY 1574 ATCTTGCCCTTTTGTACATGAGGTCACCTCTGAACTGATTTTTTTTTTGTAGTAAAGAAAT 1633  
Db 1721 ATCTTGCCCTTTTGTACATGAGGTCACCTCTGAACTGATTTTTTTTTTGTAGTAAAGAAAT 1780  
QY 1634 CTGTGACTATTTATTTCTTAAAGGAATA 1662

Db 1781 CTTGACTACTTTTAAAAA 1809

RESULT 15

US-60-474-733-338  
; Sequence 338, Application US/60474733  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Isfort, Robert J  
; TITLE OF INVENTION: Methods for Identifying Compounds for Regulating Muscle Mass or  
; TITLE OF INVENTION: Function  
; FILE REFERENCE: 92599  
; CURRENT APPLICATION NUMBER: US/60/474,733  
; CURRENT FILING DATE: 2003-05-30  
; NUMBER OF SEQ ID NOS: 353  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 338  
; LENGTH: 1814  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1446)  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1562)..(1562)  
; OTHER INFORMATION: n is a, c, g, or t  
US-60-474-733-338

Query Match 72.3%; Score 1632; DB 117; Length 1814;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 14 GTCAATTTTAAAGCAAGGTCCTTTGAATGACGAGATATTCGGACCGGAGATACGTTG 73  
Db 161 GTCAATTTTAAAGCAAGGTCCTTTGAATGACGAGATATTCGGACCGGAGATACGTTG 220  
QY 74 ACGAATACAGGAATGACTTCTGTGAAGGATATGTTCTCTAGACATTTATCACAGACATTTG 133  
Db 221 ACGAATACAGGAATGACTTCTGTGAAGGATATGTTCTCTAGACATTTATCACAGACATTTG 280  
QY 134 AAAGCGGATTCGAATCCATGTCGAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCCCTA 193  
Db 281 AAAGCGGATTCGAATCCATGTCGAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCCCTA 340  
QY 194 AAAGCAAGCGCAATAGACACATGTTTCAAGTCATCAGTCACGTTTCAAGAGCACCAGAAAGGA 253  
Db 341 AAAGCAAGCGCAATAGACACATGTTTCAAGTCATCAGTCACGTTTCAAGAGCACCAGAAAGGA 400  
QY 254 AAAGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACG 313  
Db 401 AAAGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACG 460  
QY 314 TTCTAAGCAAGATATGAAATCGTGGACATCTTTGGGTGAAGGAGCCTTTGGCAAGTTG 373  
Db 461 TTCTAAGCAAGATATGAAATCGTGGACATCTTTGGGTGAAGGAGCCTTTGGCAAGTTG 520  
QY 374 TAGAGTGCATTTGATCATGCGCATGATGCGCATGTCATGTCAGTGAAGATTCGTAAATAATG 433  
Db 521 TAGAGTGCATTTGATCATGCGCATGATGCGCATGTCATGTCAGTGAAGATTCGTAAATAATG 580  
QY 434 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAAGACATTTAAATAGTA 493  
Db 581 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAAGACATTTAAATAGTA 640  
QY 494 CTGATCCCAATAGTGTCTTCGATGTCGAGATGTCGAGATGTCGATCATCATGTC 553  
Db 641 CTGATCCCAATAGTGTCTTCGATGTCGAGATGTCGAGATGTCGATCATCATGTC 700  
QY 554 ATGTTTGTATTTGTTTGAACCTTACTGGGACCTTAGTACTTACGATTTTCAATTAAGAAACA 613  
Db 701 ATGTTTGTATTTGTTTGAACCTTACTGGGACCTTAGTACTTACGATTTTCAATTAAGAAACA 760





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 07:34:52 ; Search time 7297.99 Seconds  
(without alignments)

11766.665 Million cell updates/sec

Title: US-10-801-671-1\_COPY\_72\_2327

Perfect score: 2256

Sequence: 1 gacactcaccagtcattt.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gse1:\*

9: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description         |
|------------|--------|-------------|--------|-------|---------------------|
| 1          | 1644.4 | 72.9        | 2798   | 3     | CR749504            |
| 2          | 1283.2 | 56.9        | 1556   | 3     | BC015942 Homo sapi  |
| 3          | 1194   | 52.9        | 1446   | 9     | AY408252 Homo sapi  |
| 4          | 1118.8 | 49.6        | 1865   | 3     | AF212224 Homo sapi  |
| 5          | 1078.2 | 47.8        | 1629   | 3     | AK013974 Mus muscu  |
| 6          | 978.2  | 43.4        | 1421   | 9     | AY408254 Mus muscu  |
| 7          | 937.6  | 41.6        | 1446   | 9     | AY408253 Pan trogl  |
| 8          | 814.8  | 36.1        | 1737   | 3     | CR591789 full-leng  |
| 9          | 814.8  | 36.1        | 1746   | 3     | CR612115 full-leng  |
| 10         | 814.8  | 36.1        | 1754   | 3     | CR617472 full-leng  |
| 11         | 754.4  | 33.4        | 973    | 5     | BUS16568 AGENCOURT  |
| 12         | 734.2  | 32.5        | 1689   | 3     | AK076199 Mus muscu  |
| 13         | 722    | 32.0        | 722    | 4     | BM786882 K-EST0065  |
| 14         | 720.8  | 32.0        | 966    | 7     | CF110608 Shultzomi  |
| 15         | 700.8  | 31.1        | 727    | 1     | AA631990 np74a05.s  |
| 16         | 700.6  | 31.1        | 729    | 4     | BM675318 UI-E-EJO-  |
| 17         | 670.6  | 29.7        | 695    | 5     | BQ774390 UI-H-E21-  |
| 18         | 666    | 28.5        | 1664   | 3     | CR593610 full-leng  |
| 19         | 647.6  | 28.7        | 1775   | 3     | BC028149 Homo sapi  |
| 20         | 640    | 28.4        | 3226   | 3     | BC028573 Homo sapi  |
| 21         | 635.4  | 28.2        | 721    | 5     | BUI184090 AGENCOURT |
| 22         | 635.2  | 28.2        | 919    | 4     | BG109862 60279528   |
| 23         | 633.6  | 28.1        | 659    | 4     | BM681648 UI-E-EJO-  |
| 24         | 630    | 27.9        | 664    | 5     | BM969890 UI-CF-DU1  |

|   |    |       |      |      |   |                    |
|---|----|-------|------|------|---|--------------------|
| c | 25 | 628   | 27.8 | 1337 | 7 | CO389096           |
|   | 26 | 623.8 | 27.7 | 643  | 6 | CB127989 K-EST0177 |
|   | 27 | 621.8 | 27.6 | 849  | 4 | BG916444 602814061 |
| c | 28 | 620.8 | 27.5 | 637  | 5 | BUT29717 UI-E-CK1- |
|   | 29 | 620.6 | 27.5 | 667  | 1 | AJ656344 AJ656344  |
|   | 30 | 617   | 27.3 | 1551 | 7 | CF110517 Shultzomi |
|   | 31 | 611.6 | 27.1 | 957  | 5 | BX374503 BX374503  |
|   | 32 | 609.4 | 27.0 | 923  | 5 | BX385193 BX385193  |
| c | 33 | 609   | 27.0 | 767  | 6 | BX374502 BX374502  |
|   | 34 | 607   | 26.9 | 607  | 6 | CD676847 ho07d04.y |
|   | 35 | 594.8 | 26.4 | 1138 | 3 | CR609176 full-leng |
|   | 36 | 592.2 | 26.2 | 902  | 1 | AL540351 AL540351  |
|   | 37 | 591.6 | 26.2 | 950  | 7 | CF411264 CH3#072.F |
| c | 38 | 591   | 26.2 | 624  | 4 | BI494554 df11a06.  |
|   | 39 | 589.8 | 26.1 | 893  | 6 | CB202758 AGENCOURT |
|   | 40 | 589.2 | 26.1 | 950  | 5 | BP383242 BP383242  |
| c | 41 | 578.4 | 25.6 | 593  | 6 | CA314357 UI-CF-PNO |
|   | 42 | 573.4 | 25.4 | 710  | 5 | BQ210018 UI-R-EPO- |
| c | 43 | 570.8 | 25.3 | 618  | 6 | CB285812 CMD35.A05 |
|   | 44 | 570   | 25.3 | 581  | 5 | BP307701 BP307701  |
|   | 45 | 566   | 25.1 | 568  | 4 | BM727179 UI-E-EJO- |

## ALIGNMENTS

RESULT 1  
CR749504 2798 bp mRNA linear HTC 19-AUG-2004  
LOCUS Homo sapiens mRNA; cDNA DKFZp686A20267 (from clone DKFZp686A20267).  
DEFINITION CR749504  
ACCESSION CR749504  
VERSION CR749504.1 GI:51476700  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2798)  
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.  
CONSRMT The German cDNA Consortium  
TITLE Direct Submission  
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Dueseldorff/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686A20267) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering.  
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686A20267  
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.  
FEATURES  
source  
1..2798  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686A20267"  
/tissue type="cerebellum"  
/clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host DH10B; sites SfiI + SfiIb"  
/dev\_stages="adult"  
/note="protein serine threonine kinase Ctk4, differentially spliced"  
1..2798  
/gene="DKFZp686A20267"  
85..246  
/gene="DKFZp686A20267"  
/codon\_start=1  
/product="hypothetical protein"

/protein id="CAH18325.1"  
/db\_xref="GI:51476701"  
/translation="WRHSKRTHCPDWDSRESWGHSYRGSHKRRSHSSSQENRHCK  
PHHQFKESD"

ORIGIN

|                       |                 |  |           |              |  |
|-----------------------|-----------------|--|-----------|--------------|--|
| Query Match           | 72.9%;          | Score 1644.4;  | DB 3;     | Length 2798; |  |
| Best Local Similarity | 99.3%;          | Pred. No. 0;   |           |              |  |
| Matches 1651;         | Conservative 0; | Mismatches 11;   | Indels 0; | Gaps 0;      |  |
| QY                    | 1               | GACACTCATCCGAGTCATATTTAGAACCAAGGTCCTTGAATGAGCGAGATATCGGGAC     | 60        |              |  |
| DB                    | 1135            | GACACTCATCCGAGTCATATTTAGAACCAAGGTCCTTGAATGAGCGAGATATCGGGAC     | 1194      |              |  |
| QY                    | 61              | CGGAGATACGTTGACGAATACAGGAATCACTACTCTGTGAAGGATATGTTCCCTAGACATAT | 120       |              |  |
| DB                    | 1195            | CGGAGATACGTTGACGAATACAGGAATCACTACTCTGTGAAGGATATGTTCCCTAGACATAT | 1254      |              |  |
| QY                    | 121             | CACAGAGACATTTGAAGCGGGTATCGAATCCACTGCGAGTAAATCTTCACTCCGACGACGG  | 180       |              |  |
| DB                    | 1255            | CACAGAGACATTTGAAGCGGGTATCGAATCCACTGCGAGTAAATCTTCACTCCGACGACGG  | 1314      |              |  |
| QY                    | 181             | AGAAGCAGTCTTAAAGAGCGGCATAGACACTGTTCAAGTCATCAGTCAGTTCGAA        | 240       |              |  |
| DB                    | 1315            | AGAAGCAGTCTTAAAGAGCGGCATAGACACTGTTCAAGTCATCAGTCAGTTCGAA        | 1374      |              |  |
| QY                    | 241             | AGCCACCGAAGAAAGATCAGAGATATAGAGGATGATGAGGAGGGTCACCTGATCTGT      | 300       |              |  |
| DB                    | 1375            | AGCCACCGAAGAAAGATCAGAGATATAGAGGATGATGAGGAGGGTCACCTGATCTGT      | 1434      |              |  |
| QY                    | 301             | CAAAAGTGAGAGCGTTCTTAAGAGCAAGATATGAATFCGTGGACACTTTGGGTGAAGGACC  | 360       |              |  |
| DB                    | 1435            | CAAAAGTGAGAGCGTTCTTAAGAGCAAGATATGAATFCGTGGACACTTTGGGTGAAGGACC  | 1494      |              |  |
| QY                    | 361             | TTTGGCAAAAGTTGATAGTGCATGATCATGCGATGGATGCGATGCGATGAGAGTGA       | 420       |              |  |
| DB                    | 1495            | TTTGGCAAAAGTTGATAGTGCATGATCATGCGATGGATGCGATGCGATGAGAGTGA       | 1554      |              |  |
| QY                    | 421             | ATCGTAAAGAAATGTAGGCGGTTTACCGTGAAGCAGCTCGTTGAGAAATCCAAAGTATTAG  | 480       |              |  |
| DB                    | 1555            | ATCGTAAAGAAATGTAGGCGGTTTACCGTGAAGCAGCTCGTTGAGAAATCCAAAGTATTAG  | 1614      |              |  |
| QY                    | 481             | CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGATGGTTT     | 540       |              |  |
| DB                    | 1615            | CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGATGGTTT     | 1674      |              |  |
| QY                    | 541             | GATCATATGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG    | 600       |              |  |
| DB                    | 1675            | GATCATATGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG    | 1734      |              |  |
| QY                    | 601             | ATTAAGAAACACAGCTTTCTGCAATTTCAAAATGACCAATCAGGCAGATGCGGTATCAG    | 660       |              |  |
| DB                    | 1735            | ATTAAGAAACACAGCTTTCTGCCATTTCAATTTGACCAATCAGGCAGATGCGGTATCAG    | 1794      |              |  |
| QY                    | 661             | ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT    | 720       |              |  |
| DB                    | 1795            | ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT    | 1854      |              |  |
| QY                    | 721             | GAAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG     | 780       |              |  |
| DB                    | 1855            | GAAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG     | 1914      |              |  |
| QY                    | 781             | GATGAACGCACACTGAAACACACAGATATCAAAATGTTGTTGTTGTTGTTGTTGTTGTTG   | 840       |              |  |
| DB                    | 1915            | GATGAACGCACACTGAAACACACAGATATCAAAATGTTGTTGTTGTTGTTGTTGTTGTTG   | 1974      |              |  |
| QY                    | 841             | GATGATGAACATCACAGTACTTGTGTCTACCCGGCATACAGAGCTCCCGAGTCAAT       | 900       |              |  |
| DB                    | 1975            | GATGATGAACATCACAGTACTTGTGTCTACCCGGCATACAGAGCTCCCGAGTCAAT       | 2034      |              |  |
| QY                    | 901             | TTGGCTTTAGTGTGCTCAGGCTGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG     | 960       |              |  |
| DB                    | 2035            | TTGGCTTTAGTGTGCTCAGGCTGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG     | 2094      |              |  |

|    |      |   |      |  |  |
|----|------|---|------|--|--|
| QY | 961  | TATTACCTTGGTTTTCACAGTCTTTTACAGCTCATGATAGTAAAGAGCACCTTGGCAATGATG | 1020 |  |  |
| DB | 2095 | TATTACCTTGGTTTTCACAGTCTTTTACAGCTCATGATAGTAAAGAGCACCTTGGCAATGATG | 2154 |  |  |
| QY | 1021 | GAACGAATATTAGGACCCATACCAACACATGATTGAGAAAACAGAAAACGCAAGTAT       | 1080 |  |  |
| DB | 2155 | GAACGAATATTAGGACCCATACCAACACATGATTGAGAAAACAGAAAACGCAAGTAT       | 2214 |  |  |
| QY | 1081 | TTTCACCATACACAGCTAGATTGGATGAACACACAGTTCTGCTGCTAGATATGTTAGGAGA   | 1140 |  |  |
| DB | 2215 | TTTCACCATACACAGCTAGATTGGATGAACACACAGTTCTGCTGCTAGATATGTTAGGAGA   | 2274 |  |  |
| QY | 1141 | CGCTGCAAAACCGTTTGAAGGAATTTATGCTTTTGTCTCATGATGAAGAACATGAGAAA     | 1200 |  |  |
| DB | 2275 | CGCTGCAAAACCGTTTGAAGGAATTTATGCTTTTGTCTCATGATGAAGAACATGAGAAA     | 2334 |  |  |
| QY | 1201 | GACCTGGTTCGAAGAATGTTAGATATGATCAACTCAAGAATTTACCTTCGGATGAAGCA     | 1260 |  |  |
| DB | 2335 | GACCTGGTTCGAAGAATGTTAGATATGATCAACTCAAGAATTTACCTTCGGATGAAGCA     | 2394 |  |  |
| QY | 1261 | TTGCAGCATCTCTTTCTTTGACTTATTAAGAAAGAAATGAAATGGAAATGAGTGGTCTTAC   | 1320 |  |  |
| DB | 2395 | TTGCAGCATCTCTTTCTTTGACTTATTAAGAAAGAAATGAAATGGAAATGAGTGGTCTTAC   | 2454 |  |  |
| QY | 1321 | TATATACCTCTCTAGAGAGATTACTTTAAGACTGTGTGAGTCAACTAAACATTTCTAATAT   | 1380 |  |  |
| DB | 2455 | TATATACCTCTCTAGAGAGATTACTTTAAGACTGTGTGAGTCAACTAAACATTTCTAATAT   | 2514 |  |  |
| QY | 1381 | TTTTGTAACATTAATTTTGTACAGTTTAAAGTGAATTAATTTGTTGTTGTTGTTGTTGTTG   | 1440 |  |  |
| DB | 2515 | TTTTGTAACATTAATTTTGTACAGTTTAAAGTGAATTAATTTGTTGTTGTTGTTGTTGTTG   | 2574 |  |  |
| QY | 1441 | TAGCATTAATTAACCTTTGTTAAGCAAGTATGGTCTTGAATTAATGCAATAGAAAAATTA    | 1500 |  |  |
| DB | 2575 | TAGCATTAATTAACCTTTGTTAAGCAAGTATGGTCTTGAATTAATGCAATAGAAAAATTA    | 2634 |  |  |
| QY | 1501 | AATTTTCTTTTTCGAAATACCATTTTAAATACCTTTGAAATATCCCTTTGTTGTTGTTG     | 1560 |  |  |
| DB | 2635 | AATTTTCTTTTTCGAAATACCATTTTAAATACCTTTGAAATATCCCTTTGTTGTTGTTG     | 2694 |  |  |
| QY | 1561 | ATAAATGTTGATGTTGCTGCTTTTGTACATGGAGTCACTCTGCACTGATTTTTTTTGG      | 1620 |  |  |
| DB | 2695 | ATAAATGTTGATGTTGCTGCTTTTGTACATGGAGTCACTCTGCACTGATTTTTTTTGG      | 2754 |  |  |
| QY | 1621 | AGTAAAGGAAATCTTGACTACTTTATATATCTTTAAAGGAATA                     | 1662 |  |  |
| DB | 2755 | AGTAAAGGAAATCTTGACTACTTTATATATCTTTAAAGGAATA                     | 2796 |  |  |

|            |  |             |      |        |                 |
|------------|--|-------------|------|--------|-----------------|
| RESULT 2   | BC015942   | 1556 bp     | mRNA | linear | HTC 19-NOV-2003 |
| LOCUS      | BC015942   |             |      |        |                 |
| DEFINITION | Homo sapiens CDC-like kinase 4, mRNA (cdna clone IMAGE:3908500), containing frame-shift errors.  |             |      |        |                 |
| ACCESSION  | BC015942   |             |      |        |                 |
| VERSION    | BC015942.1   | GI:16198521 |      |        |                 |
| KEYWORDS   | HTC.   |             |      |        |                 |
| SOURCE     | Homo sapiens (human)   |             |      |        |                 |
| ORGANISM   | Homo sapiens   |             |      |        |                 |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  |             |      |        |                 |
| REFERENCE  | 1 (bases 1 to 1556)  |             |      |        |                 |
| AUTHORS    | Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, C.M., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., |             |      |        |                 |

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
12477932  
2 (bases 1 to 1556)  
Strausberg, R.  
Direct Submission  
Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Boedet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carlie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 15 Row: d Column: 16  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10190705  
This clone has the following problem: frame shifted.

Location/Qualifiers  
1..1556  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3908500"  
/tissue\_type="Uterus, leiomyosarcoma"  
/clone\_lib="NIH MGC\_71"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

ORIGIN  
Query Match 56.9%; Score 1283.2; DB 3; Length 1556;  
Best Local Similarity 99.8%; Pred No. 2e-265;  
Matches 1285; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 14 GTCATTATTAGAACGAGGCTCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 73  
Db 269 GTCATTATTAGAACGAGGCTCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 328  
Qy 74 ACGAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTATCACAGACATTG 133  
Db 329 ACGAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTATCACAGACATTG 388  
Qy 134 AAAGCGGTATCGAATCCACATGCGAGTAATCTTCAGTCCGACGAGGAGGAGCGAGTCCTA 193  
Db 389 AAAGCGGTATCGAATCCACATGCGAGTAATCTTCAGTCCGACGAGGAGGAGCGAGTCCTA 448  
Qy 194 AAAGGAAGCGCAATAGACACATGTTCAAGTCATCAGTCACGTTTCGAAGACCCACCGAAGGA 253

Db 449 AAAGGAAGCGCAATAGACACATGTTCAAGTCATCAGTCACGTTTCGAAGACCCACCGAAGGA 508  
Qy 254 AAAGATCCAGAGATATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGACG 313  
Db 509 AAAGATCCAGAGATATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGACG 568  
Qy 314 TTCTAAGACGAAGATATGAAATCGTGGACACATTTGGGTGAAGGAGCCCTTTGGCAAAGTTG 373  
Db 569 TTCTAAGACGAAGATATGAAATCGTGGACACATTTGGGTGAAGGAGCCCTTTGGCAAAGTTG 628  
Qy 374 TAGAGTCATTTGATCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCAT 433  
Db 629 TAGAGTCATTTGATCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCAT 688  
Qy 434 TAGCGCGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 493  
Db 689 TAGCGCGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 748  
Qy 494 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGAAATGGTTTGCATCATCATGTC 553  
Db 749 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGAAATGGTTTGCATCATCATGTC 808  
Qy 554 ATGTTTGTATTTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 613  
Db 809 ATGTTTGTATTTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 868  
Qy 614 GCTTTTCGCCATTTCAAATTCAGCACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 673  
Db 869 GCTTTTCGCCATTTCAAATTCAGCACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 928  
Qy 674 TAAATTTTACATATAATAAATTAACCCATACAGATCTCAAGCCTCAAAATATTTTGT 733  
Db 929 TAAATTTTACATATAATAAATTAACCCATACAGATCTCAAGCCTCAAAATATTTTGT 988  
Qy 734 TTGTGAAGTCGACTATGTAGTCAAAATATAAATTTCAAATGAACGCTGATGACGACAC 793  
Db 989 TTGTGAAGTCGACTATGTAGTCAAAATATAAATTTCAAATGAACGCTGATGACGACAC 1048  
Qy 794 TGAAGAACACAGATATCAAGTCTGTGACTTTTGAAGTGAACGCTGATGATGATGATGATGAT 853  
Db 1049 TGAAGAACACAGATATCAAGTCTGTGACTTTTGAAGTGAACGCTGATGATGATGATGATGAT 1108  
Qy 854 ACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTCCCGAGTCAATTTTGGCTTTAGTT 913  
Db 1109 ACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTCCCGAGTCAATTTTGGCTTTAGTT 1168  
Qy 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGTTTGCATTTCTTTTGAATATTACCTTGGTT 973  
Db 1169 GGTCTCAGCCTTTGATGTTTGGAGCATAGTTTGCATTTCTTTTGAATATTACCTTGGTT 1228  
Qy 974 TCAGAGTCTTTACAGCTCATGATAGTAAGAGCAGCTGGCAATGATGGAACGAATATTAG 1033  
Db 1229 TCAGAGTCTTTACAGCTCATGATAGTAAGAGCAGCTGGCAATGATGGAACGAATATTAG 1288  
Qy 1034 GACCCATACCAACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCACATAACC 1093  
Db 1289 GACCCATACCAACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCACATAACC 1348  
Qy 1094 AGCTAGATTGGGATGAACACAGTTTCTGCTGGTAGATATGTTTAGGAGACGCTGCAACCGT 1153  
Db 1349 AGCTAGATTGGGATGAACACAGTTTCTGCTGGTAGATATGTTTAGGAGACGCTGCAACCGT 1408  
Qy 1154 TGAAGGAATTTATGCTTTGTCATGATGAGAACATGAGAAATCTTTGACCTGGTTCGAA 1213  
Db 1409 TGAAGGAATTTATGCTTTGTCATGATGAGAACATGAGAAATCTTTGACCTGGTTCGAA 1468  
Qy 1214 GAATGTTAGGAATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGCATTTGACGATCCTT 1273  
Db 1469 GAATGTTAGGAATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGCATTTGACGATCCTT 1528  
Qy 1274 TCCTTTGACTTATTAAAAAGAAATGAAA 1301

Db 1529 TCTTTGACTTATTAAAAA 1556

RESULT 3  
LOCUS AY408252 1446 bp DNA linear GSS 15-DEC-2003  
DEFINITION Homo sapiens CLK4 gene, VIRTUAL TRANSCRIPT, partial sequence,  
AY408252 genomic survey sequence.  
ACCESSION AY408252  
VERSION AY408252.1 GI:39764223  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 1446)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source location/Qualifiers  
1..1446  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>1446  
/gene="CLK4"  
/locus\_tag="HCW3162"

ORIGIN  
Query Match 52.9%; Score 1194; DB 9; Length 1446;  
Best Local Similarity 92.8%; Pred No. 3.2e-246;  
Matches 1194; Conservative 0; Mismatches 92; Indels 0; Gaps 0;  
Qy 14 GTCATTATTTAGAACGAGGTCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG 73  
Db 161 GTCATTATTTAGAACGAGGTCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG 220  
Qy 74 ACGAATACAGGAATGACTCTGTGAGGATATGTTCTTAGACATTATCAGAGACATTG 133  
Db 221 ACGAATACAGGAATGACTCTGTGAGGATATGTTCTTAGACATTATCAGAGACATTG 280  
Qy 134 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCCTA 193  
Db 281 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCCTA 340  
Qy 194 AAAGGACGCAATAGACACATGTTTCAAGTCATCAGTCACTGTTGAGAGACCGACGAGGA 253  
Db 341 AAAGGACGCAATAGACACATGTTTCAAGTCATCAGTCACTGTTGAGAGACCGACGAGGA 400  
Qy 254 AAAGTCCAGAGTATAGAGATGATGAGGGGTCACTGATCTGTCAAAAGTGAGACG 313  
Db 401 AAAGTCCAGAGTATAGAGATGATGAGGGGTCACTGATCTGTCAAAAGTGAGACG 460  
Qy 314 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTGGCAAAAGTTG 373  
Db 461 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTGGCAAAAGTTG 520  
Qy 374 TAGAGTGCAATGATCATGGCATGGATGGCATGTCATGTAGCAGTGAAATCGTAAAAAATG 433

Db 521 NNNNNNNNNNNNNNNNNNNNGATGGCATGTCATGTAGCAGTGAAATCGTAAAAAATG 580  
Qy 434 TAGGCCGTTACCGTGAAGCAGCTCGTTTCAAGAAATCCAAAGTATTAGAGCACCTTAAATAGTA 493  
Db 581 TAGGCCGTTACCGTGAAGCAGCTCGTTTCAAGAAATCCAAAGTATTAGAGCACCTTAAATAGTA 640  
Qy 494 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGAAATGTTTGTGATCATCATGTC 553  
Db 641 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGAAATGTTTGTGATCATCATGTC 700  
Qy 554 ATGTTTGTATGTTGTTGAACCTAGTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 613  
Db 701 ATGTTTGTATGTTGTTGAACCTAGTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 760  
Qy 614 GCTTCTGCGCAATTTCAAAATTGACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 673  
Db 761 GCTTCTGCGCAATTTCAAAATTGACACATCAGGAGATGGCGTATGGCGNNNNNNNNNNNNNN 820  
Qy 674 TAAATTTTTTACATCATATAAATAAACCCATACAGATCTGAAGCCCTGAAAAATATTTGT 733  
Db 821 NNNNNNTTTTACATCATATAAATAAACCCATACAGATCTGAAGCCCTGAAAAATATTTGT 880  
Qy 734 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAATGAAACGTGATGAACGCAC 793  
Db 881 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAATGAAACGTGATGAACGCAC 940  
Qy 794 TGAATAAACACAGATATCAAAGTGTCTGACTTTTGAAGTGCACCTGATGATGATGAACATC 853  
Db 941 TGAATAAACACAGATATCAAAGTGTCTGACTTTTGAAGTGCACCTGATGATGATGAACATC 1000  
Qy 854 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCAATTTGGCTTTAGGTT 913  
Db 1001 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCAATTTGGCTTTAGGTT 1060  
Qy 914 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTTGTCATCTTATTGAATATTACCTTGGTT 973  
Db 1061 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTTGTCATCTTATTGAATATTACCTTGGTT 1120  
Qy 974 TCAGAGTCTTTTCAAGTCTCATGATAGTAAAGAGCAGCTGGCAATGATGGAACGAATATTAG 1033  
Db 1121 TCAGAGTCTTTTCAAGTCTCATGATAGTAAAGAGCAGCTGGCAATGATGGAACGAATATTAG 1180  
Qy 1034 GACCCATACCAACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCCAATAACC 1093  
Db 1181 GACCCATACCAACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCCAATAACC 1240  
Qy 1094 AGCTAGATTGGGATGAACACAGTTCTGCTGTAGATATGTTAGGAGCGCTGCAACCGT 1153  
Db 1241 AGCTAGATTGGGATGAACACAGTTCTGCTGTAGATATGTTAGGAGCGCTGCAACCGT 1300  
Qy 1154 TGAAGGAATTTATGCTTTGTTCATGATGAAGAAACATGAGAAACTGTTGACCTGGTTCGAA 1213  
Db 1301 TGAAGGAATTTATGCTTTGTTCATGATGAAGAAACATGAGAAACTGTTGACCTGGTTCGAA 1360  
Qy 1214 GAATTTAGAAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATGTGAGCATCCTT 1273  
Db 1361 GAATTTAGAAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATGTGAGCATCCTT 1420  
Qy 1274 TCTTTGACTTATTAAGAAAGAAATGA 1299  
Db 1421 TCTTTGACTTATTAAGAAAGAAATGA 1446

RESULT 4  
AF212224  
LOCUS AF212224  
DEFINITION Homo sapiens CLK4 mRNA, complete cds.  
ACCESSION AF212224  
VERSION AF212224.1 GI:9437514  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens







```

QY 733 TTGTGAAGTCTGACTATGTAGTCAAAATATATAATCTAAATGAAACGTCATGAACGCACA 792
Db |||
QY 713 TTGTGAAGTCTGACTATGTAGTCAAAATACAAATCTAAATGAAACGAGATGAGCGACA 772
Db |||
QY 793 CTGAAGAAACACAGATATCAAAAGTTGTCATTTGGAAGTGAACATATATGATGAACAT 852
Db |||
QY 773 TTGAAGAAACACAGATATCAAAAGTTGTCATTTGGAAGTGAACATATATGATGAACAT 832
Db |||
QY 853 CACAGTACTTTTGGTCTCTACCCGCGCACTACAGAGCTCCGAGGTCAATTTGGCTTTAGGT 912
Db |||
QY 833 CATAGTACTTTTGGTCTCTACCAAGGCACTACAAGGCTCCAGAGTCAATTTGGCTTTAGGT 892
Db |||
QY 913 TGGTCTCAGCCTTGTGATGTTTGGAGCATATGATGTCATCTTATTAATTAATTAATTAAT 972
Db |||
QY 893 TGGTCTCAGCCTTGTGATGTTTGGAGCATATGATGTCATCTTATTAATTAATTAATTAAT 952
Db |||
QY 973 TTCACAGTCTTTTACAGATCATGATAGTAAGAGCACTTGGCAATGATGAGCAATATTA 1032
Db |||
QY 953 TTCACAGTCTTTTACAGATCATGATAGTAAGAGCACTTGGCAATGATGAGCAATATTA 1012
Db |||
QY 1033 GGACCCATACCAACACATGATTCAGAAACCAAGAAACGCAAGTATTTTCAACATAAC 1092
Db |||
QY 1013 GGACCCATACCAACACATGATTCAGAAACCAAGAAACGCAAGTATTTTCAACATAAC 1072
Db |||
QY 1093 CAGCTAGATTTGGATGAACACAGTTCTCTGCTGATAGATATGTTAGAGACGCTGCAAAACCG 1152
Db |||
QY 1073 CAGCTAGATTTGGATGAACACAGTTCTCTGCTGATAGATATGTTAGAGACGCTGCAAAACCG 1132
Db |||
QY 1153 TTGAAGGAATTTATGCTTTGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
Db |||
QY 1133 TTGAAGGAATTTATGCTTTGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1192
Db |||
QY 1213 AGAATGTTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
Db |||
QY 1193 AGAATGTTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1252
Db |||
QY 1273 TTCTTTGACTTATTAAGAAAGAAATGAG-TGGAGTCAAGGCTCTT-CTGTGATCTTCTC 1332
Db |||
QY 1333 TAGAAGAGATTTACTT-AGACGTGTGTCAGTCAATTAACATCTTAATTTTGTGAACA 1391
Db |||
QY 1311 TAGGAGAGATTTACTTCCAGACTGTGTCAGTCAATTAACATCTTAATTTTGTGAACA 1370
Db |||
QY 1392 TTAATATTTTGTACAGTTAAGTCTAAATTTGATGATGATGATGATGATGATGATGATGATGAT 1451
Db |||
QY 1371 TTAATATTTTGTACAGTTAAGTCTAAATTTGATGATGATGATGATGATGATGATGATGATGAT 1430
Db |||
QY 1452 ACTTGTTAAGCAAGTATGCTCTTGATAATGCAATTAAGAAATTAATTTTCTTTT 1511
Db |||
QY 1431 CTTTGTTAAGTATGCTGTGATAATGCAATTAAGCAAGTATTAATTTTCTTTT 1490
Db |||
QY 1512 T--TGAATTAACATTTTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATATGTG 1569
Db |||
QY 1491 TGAATTTAATTTGCCACTTTAAGGGCTTTTGAATGCCCCTTTTGTGTCAGTGAATATGTG 1550
Db |||
QY 1570 ATTGATCTTGTCTTTGACATGAGGTCACTCTGGAAGTCAATTTTCTTTT-----GAG 1622
Db |||
QY 1551 ATTGATCTTGTCTTTGACATGAGGTCACTCTGGAAGTCAATTTTCTTTT-----GAG 1600
Db |||
QY 1623 TAAAGGAAATCTTTGAC 1639
Db |||
QY 1611 TAAAGGAAATCTTTGAC 1627
Db |||

```

```

RESULT 6
AY408254
LOCUS AY408254
DEFINITION Mus musculus CLK4 gene, LINEAR TRANSCRIPT, partial sequence,
AY408254
ACCESSION AY408254
VERSION AY408254.1 GI:39764225

```

```

KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1421)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1421)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
location/Qualifiers
1..1421
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1421
/gene="CLK4"
/locus_tag="HCM3162"
gene
ORIGIN
Query Match 43.4%; Score 978.2; DB 9; Length 1421;
Best Local Similarity 84.1%; Pred. No. 9e+200;
Matches 1081; Conservative 0; Mismatches 180; Indels 25; Gaps 1;
QY 14 GTCATTATTTAGAAGCAAGTCTTTGAATGAGCAGATATTATCGGACCGGAGATACGTTG 73
Db |||
QY 161 GTCATTATTTAGAAGCAAGTCTTTGAATGAGCAGATATTATCGGACCGGAGATACGTTG 220
Db |||
QY 74 ACGAATACAGGAATGACTTCTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG 133
Db |||
QY 221 ATGAATACAGGAATGACTTCTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG 280
Db |||
QY 134 AAACGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGAGCGAGAGAGACATCTCTA 193
Db |||
QY 281 AAAGCACTTACCGGATCCATTCGCAATTAATCTTCAGTCCGAGCGAGAGAGACATCTCTA 340
Db |||
QY 194 AAAGCAAGCGCAATAGACACTGTTCAAGTCAATGATGATGATGATGATGATGATGATGATGAT 253
Db |||
QY 341 AGAAGAAAGCTAATAGACCTTGTGCAAGTCAATGATGATGATGATGATGATGATGATGATGAT 400
Db |||
QY 254 AAAGTCCAGGATATAGAGGATGATGAGGAGGTTCACCTGATCTGTCAAAAGTGGAGACG 313
Db |||
QY 401 AAAGTCCAGGATATAGAGGATGATGAGGAGGTTCACCTGATCTGTCAAAAGTGGAGACG 460
Db |||
QY 314 TTCTAAGACGAGATATGAATCTGGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 373
Db |||
QY 461 TTCTAAGACGAGATATGAATCTGGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 520
Db |||
QY 374 TAGAGTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
Db |||
QY 521 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 580
Db |||
QY 434 TAGGCCGTTTACCGTGAAGCAGCTCGTTTCAAGAAATCCAGTATTAGAGCACTTTAAATAGTA 493
Db |||
QY 581 TAGGCCGTTTACCGGAGGAGCAGCTCGTTTCAAGAAATCCAGTATTAGAGCACTTTAAACGA 640
Db |||
QY 494 CTGATCCCAATAGTGTCTTCCGATGTTCCAGATGCTAGATATGTTTGTATCATCATGTC 553
Db |||
QY 641 CTGACCCCAACAGTGTCTTCCGATGCTCCAGATGCTAGATGCTTGTGTTGATCATCATGTC 700
Db |||

```



|            |   |  |                 |
|------------|---|--|-----------------|
| QY         | 734   | TTGTGAAGTCTGACTATGTAAGTCAATATATAATTTCTTAAATGAAACGATGAACGCACAC    | 793             |
| DB         | 881   | NTNNANGTCTGACTATGTTNGTCAATATATAATTTCTTAAATGAAACGATGAACGNNNN      | 940             |
| QY         | 794   | TGAAAAACACAGATATCAAAAGTTGTTGACATTTTGGAGTGCAACGTAATGATGATGAACATC  | 853             |
| DB         | 941   | TGAAAAACACAGATATCAAAAGTTGTTGACATTTTGGAGTGCAACGTAATGATGATGAACATC  | 1000            |
| QY         | 854   | ACAGTACTCTTGGTGTCTTACCCGGCAGCTACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT  | 913             |
| DB         | 1001  | ACAGTACTCTTGGTGTCTTACCCGGCAGCTACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT  | 1060            |
| QY         | 914   | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATGTAATATTAATCTTGGTT    | 973             |
| DB         | 1061  | GGTNNCGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATGTAATATTAATCTTGGTT     | 1120            |
| QY         | 974   | TCACAGTCTTTTCAGAGCTCATGATAGTAAGAGAGCAGCCTGGCAATGATGGAACGAATATTAG | 1033            |
| DB         | 1121  | TCACAGTCTTTTCAGAGCTCATGATAGTAAGAGAGCAGCCTGGCAATGATGGAACGAATATTAG | 1180            |
| QY         | 1034  | GACCATACCACAAACACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACC   | 1093            |
| DB         | 1181  | GACCATACCACAAACACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACC   | 1240            |
| QY         | 1094  | AGCTAGATTGGGATGAACACAGTGTGCTGCTGATATGTTAGGAGACGCTGCAAAACCGT      | 1153            |
| DB         | 1241  | NN     | 1300            |
| QY         | 1154  | TGAAGGAATTATGCTTTGTGCATGATGAAGAAACATGAGAAACCTGTTTGACCTGGTTGCAA   | 1213            |
| DB         | 1301  | TGAAGNN      | 1360            |
| QY         | 1214  | GAATGTTAGAAATATGATCCAACTCAAAGAATTACCTTGGATGAAGCATTCGAGCATCCTT    | 1273            |
| DB         | 1361  | NNNT     | 1420            |
| QY         | 1274  | TCTTTGACTTTATTAATAAAGAAATGA                                      | 1299            |
| DB         | 1421  | TCTTTGACTTTATTAATAAAGAAATGA                                      | 1446            |
| RESULT 8   |   |  |                 |
| CR591789   |   |  |                 |
| LOCUS      |   |  |                 |
| DEFINITION | CR591789  | 1737 bp  | linear          |
| ACCESSION  | full-length cDNA clone CL0BA009ZE09 of Placent                      |  | HTC 21-JUL-2004 |
| VERSION    | (human).  |  |                 |
| KEYWORDS   | CR591789  |  |                 |
| SOURCE     | HTC; CNSLT cDNA.  |  |                 |
| ORGANISM   | Homo sapiens  |  |                 |
| REFERENCE  | Homo sapiens  |  |                 |
| AUTHORS    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |  |                 |
| TITLE      | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.         |  |                 |
| JOURNAL    | Li, W.B., Gruber, C., Jessee, J., and Polayes, D.                   |  |                 |
| REMARK     | 1. (bases 1 to 1737)  |  |                 |
|            | Full-length cDNA libraries and normalization                        |  |                 |
|            | Unpublished   |  |                 |
|            | Contact : feng liang Email : fliang@lifetech.com URL :              |  |                 |
|            | http://fulllength.invitrogen.com/ Invitrogen Corporation 1600       |  |                 |
|            | Faraday Avenue  |  |                 |
|            | 2 (bases 1 to 1737)   |  |                 |
| REFERENCE  | Genoscope.  |  |                 |
| AUTHORS    | Direct Submission   |  |                 |
| TITLE      | Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : |  |                 |
| JOURNAL    | BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  |  |                 |
| COMMENT    | - web : www.genoscope.cns.fr)                                       |  |                 |
|            | 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime |  |                 |
|            | end enriched, double-strand cDNA was digested with Not I and cloned |  |                 |
|            | into the Not I and EcoK V sites of the pCMVSPORT 6 vector. Library  |  |                 |
|            | was normalized. Library was constructed by Life Technologies, a     |  |                 |
|            | division of Invitrogen.   |  |                 |
| FEATURES   | Location/Qualifiers   |  |                 |



[illegible]



```

Db      726  GATATGTTAGGAGACGCTGCAAGCGTTAAAGCAATTTATGCTGTGTATGACGAGGAGC 785
Qy      1187  ATGAGAACTGTTTGACCTGGTTCGAAGAATTTAGAAATATGATCAACTCAAGAAATTA 1246
Db      786  ATGAGAGCTGTTTGACCTGGTTC-AAAATGTTGGAGTAGACCCACCGAAGAGATCA 844
Qy      1247  -CCTTGGATGAAGCAATGAGCATCCTTCTTTGATCTATTATAAAAAAGAAATGAATGGG 1305
Db      845  TCTTTGGATGAAGCAATGAGCATCCTTCTTCTGACTTTATATAAAAAAGAAATGAATGGG 904
Qy      1306  AATCAGTGGTCTTACTATATATCTCT-AGNAGAGATTACTTATACATGTTGTCAGTCAA 1364
Db      905  AGCCAGGGGGCTTCTGGGAATCCCTTAAGACAGGTTACTCCAGATGGGTGAGCCAA 964
Qy      1365  CTAACACTT 1373
Db      965  CTAACACTT 973

RESULT 12
ACCESSION AK076199 1689 bp mRNA linear HTC 03-APR-2004
VERSION AK076199.1 GI:26096662
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11076861
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 11076861
REFERENCE 6
AUTHORS

```

```

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, K., Ohno, M., Ohnato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source
1..1689
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:3830430K12"
/db_xref="taxon:10090"
/clone="3830430K12"
/sex="female"
/misc_feature
105..1556
/tissue_type="placenta and extra embryonic tissue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18 days pregnant adult"
105..1556
/note="PROTEIN KINASE CLK1 (EC 2.7.1.-) (CLK) (PROTEIN
KINASE STY) homolog [Mus musculus] (SWISSPROT|P22518,
evidence: FASTA, 95%ID, 100%length, match=1449)
putative"
1670..1675
/note="putative"
1689
/note="putative"
ORIGIN
Query Match 32.5%; Score 734.2; DB 3; Length 1689;
Best Local Similarity 74.2%; Pred. No. 3.1e-147;
Matches 957; Conservative 0; Mismatches 324; Indels 9; Gaps 2;
Qy 12 CAGTCATTATTAGAAAGCAAGGTCTCTGAATGAGCGAGATATTCGGGACCGAGATACGT 71
Db 263 CAGCTATTATCTGGAAGCAGATCCATAAATGAGAAAGCTTATCATAGTCGACGCTATGT 322
Qy 72 TGACGATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCAGAGACAT 131
Db 323 TGATGAATACAGGAATGACTACATG--GGCTACGAGCCAGGCGCATCTTATGGAGAACC 379
Qy 132 TGAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGCGACGAGAGAGAGTCC 191
Db 380 TGGAGCAGATACAGATGATGATAGCAAGTCTCTGGTAGGATGGAAGAGAGATTA 439
Qy 192 TAAAGGAAGCGCAATAG-----ACACTGTTCAAGTCATCAGTCACGTCGGAAGAGCCA 245
Db 440 ACMAAGTAACACAGGAGTCGCCACACACACTTCGCAGCACCATTCCACGCGGAGAGTCA 499
Qy 246 CCGAAGGAAGATCCAGGAGTATAGAGGATGATAGAGGGTCACTGTCTGTCGAAG 305

```



Db 500 CCGAAGAAATAAGCGAGGAGTGATAGAGGATGATGAGGAGGTCACCTGATCTGTGACAG 559  
 QY 306 TGGAGACCTTCTAAGAGCAAGATATGAATCTGTGACACTTTGGTGAAGGAGCTTTGG 365  
 Db 560 TGGAGACCTAAGTGAAGATATGATATTGTTGATATCTTTAGGTGAAGGTGCTTTGG 619  
 QY 366 CAAAGTTGTAGAGTGCATTTGATCGGCATGATGGCATGATGATGATGATGATGATGATGAT 425  
 Db 620 AGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 679  
 QY 426 AAAAAATGAGCGCTTACCGTGAAGAGAGCTGTTTCAAGAAATCCAAATGATGATGATGAT 485  
 Db 680 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739  
 QY 486 AAATAGTACTGATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545  
 Db 740 GTATACAGACAGACCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 799  
 QY 546 TCATGCTCATGTTTGTATGTTTGTGTTGAACTACTGGGACTTAGTACTTAGTACTTAGTACT 605  
 Db 800 TCGAGATCACATTTGATGTTGTTGAGCTTCTGGGCTTAGTACTTAGTACTTAGTACTTAG 859  
 QY 606 AGAAACAGCTTTGCGCATTTCAAAATGACATCAGGAGATGAGGCTATGATGATGATGAT 665  
 Db 860 GAAAAACAGTTTCTGCGCTCTCGAATGATCATATCAGGAAGATGGCATATCAAAATATG 919  
 QY 666 CCAGTCAATAAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCCCTGAAA 725  
 Db 920 CAATGCTGTAAACTTTTGGCATAGTAATAATGACATACAGACTTTGAAGCCCTGAAA 979  
 QY 726 TATTTTGTGTTGAAAGTCTGACTATGATGATCAATATAATTTCAAAATGAAACGTTGATCA 785  
 Db 980 CATCTTATTTGTGAGTCTGACTACAGAGGCTTATATCCAAATGAAACGTTGATGA 1039  
 QY 786 AGGCACATGAAAAACAAGATATCAAAAGTTGTTGACTTTGGAAGTGAAGTGAAGTGAAG 845  
 Db 1040 ACGTACTATGATAAATCCAGATATTAAGAGTGTGACTTTGGAAGTGAAGTGAAGTGAAG 1099  
 QY 846 TGAACATCACAGTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 905  
 Db 1100 TGAACATCACAGTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1159  
 QY 906 TTTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 965  
 Db 1160 CTTGCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1219  
 QY 966 CTTGCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1025  
 Db 1220 TCTTGGATTTACAGTCTTTTCCGACTCATGATAGCAGGGAACATTTAGCAATGATGGAAG 1279  
 QY 1026 AATATTAGACCCATACACACATGATTTTCAAGAAACAGAAACAGAAACAGAAACAGAA 1085  
 Db 1280 GATTTGAGACCTTACCAAGCAATGATACAGAAACAGAAACAGAAACAGAAACAGAAAC 1339  
 QY 1086 CCATAACAGCTAGATTTGGATGAACACAGTTTCTGCTGTGATATGTTAGGAGACGCTG 1145  
 Db 1340 TCATGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1399  
 QY 1146 CAAACCGTTGAAAGAAATTTATGCTTTGATGATGATGATGATGATGATGATGATGATGAT 1205  
 Db 1400 TAAACCTCTGAGGAGTTTATGCTATCTCAGGATGCCGAACATGATGATGATGATGATGAT 1459  
 QY 1206 GGTTCGAAGATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGA 1265  
 Db 1460 CATTGGGAAATGTTGGAGATGATGATCCCGCCAAAGAAATTTCTCTCAAGAAAGCCCTAA 1519  
 QY 1266 GCATCCTTTCTTTGACTTTATTTAAAAAGAA 1295  
 Db 1520 GCATCCTTTCTTTTACCCTTTAAAAAGCA 1549

LOCUS BM786882 722 bp mRNA linear EST 05-MAR-2002  
 DEFINITION K-EST0065716 S10SNU1 Homo sapiens cDNA clone S10SNU1-17-D07 5',  
 mRNA sequence.  
 ACCESSION BM786882  
 VERSION BM786882.1 GI:19135114  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (Bases 1 to 722)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 17 row: D column: 07  
 High quality sequence stop: 722.  
 FEATURES  
 Location/Qualifiers  
 1..722  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S10SNU1-17-D07"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Lymphoblast-like"  
 /cell\_line="SNU-1"  
 /lab\_host="Top10F"  
 /clone\_lib="S10SNU1"  
 /note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;  
 Site 2: XhoI; The poly (A) + RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including SfiI  
 site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized with Superscript II using SfiI  
 oligo-dT primer. After first strand synthesis, RNA was  
 degraded by NaOH treatment and cDNA was amplified by PCR  
 reaction. The PCR products were digested with SfiI and  
 cloned into DraIII- digested pME18S-FL3 vector. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library."  
 ORIGIN  
 Query Match 32.0%; Score 722; DB 4; Length 722;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-144;  
 Matches 722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 611 ACAGCTTTCTGCCATTTCAAAATGACACATCAGCAGATGCGGTATCAGATCTCCAGT 670  
 Db 1 ACAGCTTTCTGCCATTTCAAAATGACACATCAGCAGATGCGGTATCAGATCTCCAGT 60  
 QY 671 CAATAAAATTTTTTACATCATATAATAATTAACCCATACAGATCTGAAGCCTGAAATATTTT 730  
 Db 61 CAATAAAATTTTTTACATCATATAATAATTAACCCATACAGATCTGAAGCCTGAAATATTTT 120  
 QY 731 TGTTTGCAAGTCTGACTATGTAGTCAAAATATAATTTAAATTAAGAACTGATGATGAAC 790  
 Db 121 TGTTTGCAAGTCTGACTATGTAGTCAAAATATAATTTAAATTAAGAACTGATGATGAAC 180  
 QY 791 CACTGAAAAACACAGATATCAAGTTGTTGTTTGGAGTGCACAGTATGATGATGAAC 850  
 Db 181 CACTGAAAAACACAGATATCAAGTTGTTGTTTGGAGTGCACAGTATGATGATGAAC 240



```

QY 851 ATCAGAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTGGCTTTAG 910
Db 241 ATCAGAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTGGCTTTAG 300
QY 911 GTTGTCTCAGCTTGTGATGTTTGGGACATAGGTTGCATCTTATGAATATTACCTTG 970
Db 301 GTTGTCTCAGCTTGTGATGTTTGGGACATAGGTTGCATCTTATGAATATTACCTTG 360
QY 971 GTTTCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACTGGCAATGATGGAACGAATAT 1030
Db 361 GTTTCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACTGGCAATGATGGAACGAATAT 420
QY 1031 TAGGACCATACCAACACATGATTCAGAAACAAAGAAACGCAAGTATTTCACCATAT 1090
Db 421 TAGGACCATACCAACACATGATTCAGAAACAAAGAAACGCAAGTATTTCACCATAT 480
QY 1091 ACCAGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAC 1150
Db 481 ACCAGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAC 540
QY 1151 CGTTGAAGGAATTTATGCTTTGTCTATGATGAAGAACATGAGAAACTGTTTGACCTGGTTC 1210
Db 541 CGTTGAAGGAATTTATGCTTTGTCTATGATGAAGAACATGAGAAACTGTTTGACCTGGTTC 600
QY 1211 GAAGATGTTAGATATGATCAACTCAAGAAATTAACCTTGGATGAAGCATTGCAGCATC 1270
Db 601 GAAGATGTTAGATATGATCAACTCAAGAAATTAACCTTGGATGAAGCATTGCAGCATC 660
QY 1271 CTTTCTTTGACTTATTAAGAAATGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1330
Db 661 CTTTCTTTGACTTATTAAGAAATGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 720
QY 1331 TC 1332
Db 721 TC 722

RESULT 14
LOCUS CF110608
DEFINITION Shultzomica03859 Rat lung airway and parenchyma cDNA libraries
Rattus norvegicus cDNA clone Contig3421 5', mRNA sequence.
ACCESSION CF110608
VERSION CF110608.1 GI:33166640
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 966)
AUTHORS Shultz, M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Fannuchi, M.V.,
Padua, A.M., Gurske, W.A., Morin, D., Penn, S.G., Jovanovich, S.B.,
Plopper, C.G. and Buckpitt, A.R.
TITLE Gene expression analysis in response to lung toxicants: I.
JOURNAL Sequencing and microarray development
COMMENT Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 0798
Email: mashultz@ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred <
20) and vector/linker sequence has been removed.
High quality sequence stop: 966.
Location/Qualifiers
FEATURES
source
1. .966
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"

```

```

/db_xref="taxon:10116"
/clone="Contig3421"
/sex="male"
/tissue_type="airway or parenchyma"
/dev_stage="adult"
/clone_lib="Rat lung airway and parenchyma cDNA libraries"
/notes="Organ: lung; Vector: pGEM-lizf(-); Site: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."

```

## ORIGIN

```

Query Match 32.0%; Score 720.8; DB 7; Length 966;
Best Local Similarity 86.0%; Pred. No. 2.2e-144;
Matches 822; Conservative 0; Mismatches 132; Indels 2; Gaps 2;
QY 678 TTTTTCATCATTAATAAATTAACCCCATACAGATCTGAAGCTCTGAATAATTTTGTGTGT 737
Db 1 TTTTTCATCATTAATAAATTAACCCCATACAGATCTGAAGCTCTGAATAATTTTGTGTGT 60
QY 738 GAAGTCTGACTATGTAGTCAAAATATAATTTTAAATGAACGTCGATGAACGCACATGAA 797
Db 61 AAAGTCTGACTATGTAGTCAAAATATAATTTTAAATGAACGTCGATGAACGCACATGAA 120
QY 798 AAACACAGATATCAAAAGTTGTTGACTTTGGAGCTGCAACGTCATGATGATGAACATCAG 857
Db 121 AAACACAGATATCAAAAGTTGTTGACTTTGGAGCTGCAACGTCATGATGATGAACATCAG 180
QY 858 TACTTTGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTGGCTTTTGGTGTGTGTC 917
Db 181 TACTTTGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTGGCTTTTGGTGTGTGTC 240
QY 918 TCAGCTTTGTGATGTTTGGAGCATAGGTTGCATTTCTTATTGAATATTACCTTGGTTTCAC 977
Db 241 TCAGCTTTGTGATGTTTGGAGCATAGGTTGCATTTCTTATTGATTTTACCTTGGATTTCAC 300
QY 978 AGTCTTTTCAGACTCATGATAGTAAAGAGCACTCGCAATGATGGAACGAATATTAGGACC 1037
Db 301 AGTCTTTTCAGACTCATGATAGTAAAGAGCACTCGCAATGATGGAACGAATATTAGGACC 360
QY 1038 CATACCAACACATGATTTCAAGAAACAAAGAAACGCAAGTATTTTTCCACCAATACCAGCT 1097
Db 361 CATCCAGCACATGATTTCCAAAGAACAGAAACGCAAGTATTTTCCACCAATACCAGCT 420
QY 1098 AGATTGGGATGAACACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAA 1157
Db 421 AGATTGGGATGAACACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAA 480
QY 1158 GGAATTTATGCTTTGTCTATGATGAAGAACATGAGAAACTGTTTGCCTGGTTTCGAGAAAT 1217
Db 481 GGAATTTATGCTTTGTCTACCGATGAAGACATGAGAAGCTGTTTGCCTGGTTTCGAGAAAT 540
QY 1218 GTTAGAATATGATCCAACTCAAGAAATTTACCTTCGATGAAGCAATTTGACGATCTCTTCTT 1277
Db 541 GTTAGAATATGATCCAACTCAAGAAATTTACCTTCGATGAAGCAATTTGACGATCTCTTCTT 600
QY 1278 TGACTTTATTAAGAAAGAAATGAATGGGAATCAGTGGTCTTCTATATATATCTCTCTAGAA 1337
Db 601 TGACTTTATTAAGAAAGAAATGAATGGGAATCAGTGGTCTTCTATATATATCTCTCTAGAA 659
QY 1338 GAGATTAC-TTAAGACTGTGTGCTCACTCAATCAATCTTAATATTTTGTGAACATTTAA 1396
Db 660 GCAATTACTTTTAAGACTGTGTGCTCACTCAATCAATCTTAATATTTTGTGAACATTTAA 719
QY 1397 TTAATTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATATAATTAACCTTG 1456
Db 720 TTAATTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATATAATTAACCTTG 779
QY 1457 TTAAGCAAGTATGCTCTTGTATTAATGCAATGAAATAATTAATAATTTTCTTTTGTAA 1516
Db 780 TTAAGTATGTTGTTGATAATGAATGCAATGATGATGATTAATGGAATTTTCTTGTATGTTA 839
QY 1517 ATTACCAATTTTAAATACCTTTGAAATATCTTTTGTGTCAGATGATTAATGATGATGATC 1576

```

Db 840 ATTGCCATTTTAAAGGCTTTACAGTCCCTTTGTCAGTGGTAAACGCTGATTGGTC 899

QY 1577 TTGCTTTTGTACATGGAGTCACTCTGAGTGTATTTTTCAGTAAAGGAAA 1632

Db 900 TCATCTTTTGTACATTAAGGTTAACTTTGAAGTGGTTTTTTTTTCCCTCAAGTAAA 955

RESULT 15

AA631990/c

LOCUS

DEFINITION

np74a05.s1 NCI CGAP Br2 Homo sapiens cDNA clone IMAGE:1132016 3' similar to SW:KCLK\_MOUSE P22518 PROTEIN KINASE CLK ; mRNA sequence.

ACCESSION

AA631990

VERSION

AA631990.1 GI:2555404

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 727)

NCI CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapb-remail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.lnl.gov/bbrp/image/image.html](http://www-bio.lnl.gov/bbrp/image/image.html)

Insert Length: 2117 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 475.

FEATURES

source

1..727

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1132016"

/sex="female, pooled"

/tissue\_type="breast"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Br2"

/note="Vector: pVT3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pVT3 vector. This library is the normalized version of NCI CGAP Br1.1. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 31.1%; Score 700.8; DB 1; Length 727;

Best Local Similarity 99.5%; Pred. No. 4.2e-140;

Matches 724; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 917 CTCAGCCTTTGTGATTTGGAGCATAGTTGATTTCTTATTGAATATTACCTTGGTTTCA 976

Db 726 CTCAGCCTTTGTGATTTGGAGCATAGTTGGAT -CTTAAGAATATTACCTTGGTTTCA 668

QY 977 CAGTCTTTTCAGACTCATGATAGTAAGAGCAGCTGGCAATGATGGAACGAATATTAGGAC 1036

Db 667 CAGTC-TTCAGACTCATGATAGTAAGAGCAGCTGGCAATGATGGAACGAATATTAGGAC 609

QY 1037 CCATACCACACATGATTTCAGAAAACAGAAAACGCAAGTATTTTCACCATACCCAGC 1096

Db 608 CCATACCACACACATGATTTCAGAAAACAGAAAACGCAAGTATTTTCACCATACCCAGC 549

QY 1097 TAGATTGGGATGAACACAGTTCTCTGCTGTAGATATGTTAGGAGACGCTGCAGAACCGTTGA 1156

Db 548 TAGATTGGGATGAACACAGTTCTCTGCTGTAGATATGTTAGGAGACGCTGCAGAACCGTTGA 489

QY 1157 AGGAATTTATGCTTTTGTTCATGATCAAGAACATGAGAAACTGTGTTGACCTGGTTTCAAGAA 1216

Db 488 AGGAATTTATGCTTTTGTTCATGATCAAGAACATGAGAAACTGTGTTGACCTGGTTTCAAGAA 429

QY 1217 TGTAGAAATATGATCCAACTCAAGAAATTTACCTTTGGATGAGCAATTCGAGCATCCTTTCT 1276

Db 428 TGTTAGAATATGATCCAACTCAAGAAATTTACCTTTGGATGAGCAATTCGAGCATCCTTTCT 369

QY 1277 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGTCTTACTATATATCTTCTCTAGA 1336

Db 368 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGTCTTACTATATATCTTCTCTAGA 309

QY 1337 AGAGATTACTTTAAGACTGTGTCACTCAACTAAACATTTCTAATATTTTGTGTAACATTAAA 1396

Db 308 AGAGATTACTTTAAGACTGTGTCACTCAACTAAACATTTCTAATATTTTGTGTAACATTAAA 249

QY 1397 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATTAATTAACCTTG 1456

Db 248 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATTAATTAACCTTG 189

QY 1457 TTAAGCAAGTATGCTCTTGTATAATGCAATAGCAATTAAGAAATTAATTTTCTTTTGA 1516

Db 188 TTAAGCAAGTATGCTCTTGTATAATGCAATAGCAATTAAGAAATTAATTTTCTTTTGA 129

QY 1517 ATTACCAATTTTAAATACCTTTGAAATATTCCTTTGTGTCAGTGATTAATGTGATTGATC 1576

Db 128 ATTACCAATTTTAAATACCTTTGAAATATTCCTTTGTGTCAGTGATTAATGTGATTGATC 69

QY 1577 TTGCTTTTGTACATGGAGGTCACCTCTGAGTGTGATTTTTTTTGTAGTAAAGGAAATCTT 1636

Db 68 TTGCTTTTGTACATGGAGGTCACCTCTGAGTGTGATTTTTTTTGTAGTAAAGGAAATCTT 9

QY 1637 GACTACTT 1644

Db 8 GACTACTT 1

Search completed: March 16, 2005, 18:40:36

Job time : 7302.99 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 07:43:58 ; Search time 376.816 Seconds  
(without alignments)  
9796.410 Million cell updates/sec

Title: US-10-801-671-1\_COPY\_72\_2327

Perfect score: 2256  
Sequence: 1 gacactcatccagtcattt.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 2256   | 100.0       | 2354   | 3     | US-09-810-671-1    |
| 2          | 2256   | 100.0       | 2354   | 4     | US-10-109-854-1    |
| 3          | 2256   | 100.0       | 2354   | 4     | US-10-339-656-1    |
| 4          | 2046.6 | 90.7        | 2446   | 2     | US-09-016-000-9    |
| 5          | 1178.2 | 52.2        | 1456   | 4     | US-09-023-655-699  |
| 6          | 1097.4 | 48.6        | 1549   | 4     | US-09-903-999-26   |
| 7          | 1097   | 48.6        | 21234  | 3     | US-09-810-671-3    |
| 8          | 1097   | 48.6        | 21234  | 4     | US-10-109-854-3    |
| 9          | 1097   | 48.6        | 21234  | 4     | US-10-339-656-3    |
| 10         | 640.6  | 28.4        | 2254   | 4     | US-09-919-039-238  |
| 11         | 481.2  | 21.3        | 1538   | 4     | US-09-905-999-22   |
| 12         | 454.8  | 20.2        | 1787   | 4     | US-09-905-999-24   |
| 13         | 432.4  | 19.2        | 1762   | 4     | US-09-016-434-1439 |
| 14         | 430.8  | 19.1        | 1763   | 4     | US-09-949-016-2648 |
| 15         | 430.8  | 19.1        | 1763   | 4     | US-09-949-016-2649 |
| 16         | 430.8  | 19.1        | 1788   | 4     | US-09-976-594-313  |
| 17         | 263    | 11.7        | 263    | 4     | US-09-016-434-310  |
| 18         | 203.4  | 9.0         | 492    | 4     | US-09-621-976-3125 |
| 19         | 185.4  | 8.2         | 479    | 4     | US-09-621-976-3124 |
| 20         | 184.2  | 8.2         | 913    | 4     | US-09-016-434-712  |
| 21         | 123.8  | 5.5         | 378    | 1     | US-08-700-575-2    |
| 22         | 95.2   | 4.2         | 2061   | 2     | US-08-835-170-1    |
| 23         | 95.2   | 4.2         | 2061   | 3     | US-09-359-257-1    |
| 24         | 95.2   | 4.2         | 2061   | 3     | US-09-371-674-1    |
| 25         | 95.2   | 4.2         | 2327   | 2     | US-08-835-170-3    |
| 26         | 95.2   | 4.2         | 2327   | 3     | US-09-359-257-3    |
| 27         | 95.2   | 4.2         | 2327   | 3     | US-09-371-674-3    |

|    |      |     |        |   |                      |                    |
|----|------|-----|--------|---|----------------------|--------------------|
| 28 | 89   | 3.9 | 621    | 4 | US-09-248-796A-4385  | Sequence 4385, Ap  |
| 29 | 84.8 | 3.8 | 308    | 4 | US-09-621-976-9631   | Sequence 9631, Ap  |
| 30 | 81.4 | 3.6 | 260    | 4 | US-09-313-294A-2299  | Sequence 2299, Ap  |
| 31 | 75.6 | 3.4 | 3723   | 4 | US-09-949-016-3590   | Sequence 3590, Ap  |
| 32 | 75.6 | 3.4 | 71251  | 4 | US-09-949-016-15332  | Sequence 15332, A  |
| 33 | 68.8 | 3.0 | 282    | 4 | US-09-313-294A-1168  | Sequence 1168, Ap  |
| 34 | 67.2 | 3.0 | 2424   | 4 | US-09-614-221A-518   | Sequence 518, App  |
| 35 | 66.4 | 2.9 | 466    | 4 | US-09-513-999C-11373 | Sequence 11373, A  |
| 36 | 66   | 2.9 | 362    | 4 | US-09-949-016-12776  | Sequence 12776, A  |
| 37 | 63.6 | 2.8 | 187169 | 4 | US-09-949-016-15940  | Sequence 15940, A  |
| 38 | 63.6 | 2.8 | 191569 | 4 | US-09-949-016-14391  | Sequence 14391, A  |
| 39 | 61.8 | 2.7 | 25882  | 4 | US-09-949-016-14390  | Sequence 14390, A  |
| 40 | 61.8 | 2.7 | 25882  | 4 | US-09-949-016-14391  | Sequence 14391, A  |
| 41 | 60.6 | 2.7 | 1141   | 4 | US-09-806-708B-22    | Sequence 22, Appl  |
| 42 | 60.6 | 2.7 | 59519  | 4 | US-09-949-016-123504 | Sequence 123504, A |
| 43 | 60.6 | 2.7 | 119153 | 4 | US-09-949-016-12378  | Sequence 12378, A  |
| 44 | 60.4 | 2.7 | 19124  | 2 | US-08-487-826B-13    | Sequence 13, Appl  |
| 45 | 60.4 | 2.7 | 251672 | 4 | US-09-949-016-17296  | Sequence 17296, A  |

ALIGNMENTS

RESULT 1  
US-09-810-671-1  
; Sequence 1, Application US/09810671  
; Patent No. 6455291  
; GENERAL INFORMATION:  
; APPLICANT: YAN Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CLO00758  
; CURRENT APPLICATION NUMBER: US/09/810,671  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2354  
; TYPE: DNA  
; ORGANISM: Human  
US-09-810-671-1

|                       |                 |   |           |              |
|-----------------------|-----------------|---|-----------|--------------|
| Query Match           | 100.0%          | Score 2256;   | DB 3;     | Length 2354; |
| Best Local Similarity | 100.0%;         | Pred. No. 0;  |           |              |
| Matches 2256;         | Conservative 0; | Mismatches 0;   | Indels 0; | Gaps 0;      |
| Qy                    | 1               | GACACTATCCAGTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTATCGGAC       | 60        |              |
| Db                    | 72              | GACACTATCCAGTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTATCGGAC       | 131       |              |
| Qy                    | 61              | CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCCGTAGACATTAT | 120       |              |
| Db                    | 132             | CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCCGTAGACATTAT | 191       |              |
| Qy                    | 121             | CACAGAGACATTGAAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGCGACGAG   | 180       |              |
| Db                    | 192             | CACAGAGACATTGAAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGCGACGAG   | 251       |              |
| Qy                    | 181             | AGAGAGAGTCCTAAAGGAAGCGCAATAGACACTGTTCAAGTCATCATGTCAGTTCGAG    | 240       |              |
| Db                    | 252             | AGAGAGAGTCCTAAAGGAAGCGCAATAGACACTGTTCAAGTCATCATGTCAGTTCGAG    | 311       |              |
| Qy                    | 241             | AGCCACCGAAGAAAGTCCAGGATATAGAGATGATAGGAGGTCACCTGATCTGT         | 300       |              |
| Db                    | 312             | AGCCACCGAAGAAAGTCCAGGATATAGAGATGATAGGAGGTCACCTGATCTGT         | 371       |              |
| Qy                    | 301             | CAAAAGTGAGAGCGTTCTTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGGCC  | 360       |              |
| Db                    | 372             | CAAAAGTGAGAGCGTTCTTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGGCC  | 431       |              |
| Qy                    | 361             | TTTCGCAAGTTGTAGAGTCATTGATCATCGGATGGAATGATGATGATGATGATGATGATG  | 420       |              |



|                 |      |   |      |  |  |
|-----------------|------|---|------|--|--|
| ; LENGTH: 2354  |      | Query Match   |      | 100.0%; Score 2256; DB 4; Length 2354;     |  |
| ; TYPE: DNA     |      | ; ORGANISM: Homo sapien   |      | Best Local Similarity 100.0%; Pred. No. 0; |  |
| US-10-109-854-1 |      | Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;       |      |  |  |
| QY              | 1    | GACACTCATCCAGTCATTATTAGAACCAAGGTCCTTGAATGACGAGGATTTATCGGAC          | 60   |  |  |
| DB              | 72   | GACACTCATCCAGTCATTATTAGAACCAAGGTCCTTGAATGACGAGGATTTATCGGAC          | 131  |  |  |
| QY              | 61   | CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCCCTAGACATTAT       | 120  |  |  |
| DB              | 132  | CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCCCTAGACATTAT       | 191  |  |  |
| QY              | 121  | CACAGAGACNTTGAAGCGGGTATCGAATCCACTCCAGTAAATCTTCAGTCCGACGAGG          | 180  |  |  |
| DB              | 192  | CACAGAGACNTTGAAGCGGGTATCGAATCCACTCCAGTAAATCTTCAGTCCGACGAGG          | 251  |  |  |
| QY              | 181  | AGAAGCAGTCCCTAAAGGAAGGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGGAAG         | 240  |  |  |
| DB              | 252  | AGAAGCAGTCCCTAAAGGAAGGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGGAAG         | 311  |  |  |
| QY              | 241  | AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATAGGAGGGTCACCTGATCTGT           | 300  |  |  |
| DB              | 312  | AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATAGGAGGGTCACCTGATCTGT           | 371  |  |  |
| QY              | 301  | CAAAGTGAGAGCGTTCTAAGAGCAAGATATGAATCGTGGGACACTTTGGGTGAAGGAGCC        | 360  |  |  |
| DB              | 372  | CAAAGTGAGAGCGTTCTAAGAGCAAGATATGAATCGTGGGACACTTTGGGTGAAGGAGCC        | 431  |  |  |
| QY              | 361  | TTTGGCAAGTTGTAGAGTGCAATCATCATGCGATGGATGGCATGCGATGAGCAGTGAAA         | 420  |  |  |
| DB              | 432  | TTTGGCAAGTTGTAGAGTGCAATCATCATGCGATGGATGGCATGCGATGAGCAGTGAAA         | 491  |  |  |
| QY              | 421  | ATCGTAAAGAAATGTAGGCGGTTACCGGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAG       | 480  |  |  |
| DB              | 492  | ATCGTAAAGAAATGTAGGCGGTTACCGGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAG       | 551  |  |  |
| QY              | 481  | CACTTAAATAGTACATCCCAATAGTGTCTCCGATGTCCAGATGCCAGATGAGGTTT            | 540  |  |  |
| DB              | 552  | CACTTAAATAGTACATCCCAATAGTGTCTCCGATGTCCAGATGCCAGATGAGGTTT            | 611  |  |  |
| QY              | 541  | GATCATATGGTCATGTTGTATGTTGTTGAATCTTGGGACTTGTAGTACTTACCATTTTC         | 600  |  |  |
| DB              | 612  | GATCATATGGTCATGTTGTATGTTGTTGAATCTTGGGACTTGTAGTACTTACCATTTTC         | 671  |  |  |
| QY              | 601  | ATTAAGAAACACAGCTTTCTGCCATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAG       | 660  |  |  |
| DB              | 672  | ATTAAGAAACACAGCTTTCTGCCATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAG       | 731  |  |  |
| QY              | 661  | ATCTGCCAGTCAATAAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT           | 720  |  |  |
| DB              | 732  | ATCTGCCAGTCAATAAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT           | 791  |  |  |
| QY              | 721  | GAAATATTTTGTGTGAAGTCTGACTATGTAGTCAATAATAATCTTAAATGAACAGT            | 780  |  |  |
| DB              | 792  | GAAATATTTTGTGTGAAGTCTGACTATGTAGTCAATAATAATCTTAAATGAACAGT            | 851  |  |  |
| QY              | 781  | GATGAACGCACACTGAAACAAACAGATATCAAGTTGTTGACTTTTGGAAAGTGCACAGTAT       | 840  |  |  |
| DB              | 852  | GATGAACGCACACTGAAACAAACAGATATCAAGTTGTTGACTTTTGGAAAGTGCACAGTAT       | 911  |  |  |
| QY              | 841  | GATGATGAACATCAGTACTTTTGTGTCTACCCGGCACTACAGAGTCCCGAGTCAIT            | 900  |  |  |
| DB              | 912  | GATGATGAACATCAGTACTTTTGTGTCTACCCGGCACTACAGAGTCCCGAGTCAIT            | 971  |  |  |
| QY              | 901  | TTGGCTTTAGGTTGCTCTCAGCCTTGTGATGTTTCGGAGCATAGTGTGATCTTATTGAA         | 960  |  |  |
| DB              | 972  | TTGGCTTTAGGTTGCTCTCAGCCTTGTGATGTTTCGGAGCATAGTGTGATCTTATTGAA         | 1031 |  |  |
| QY              | 961  | TATTACCTTGGTTTCAAGTCTTTTTCAGACTCATGATAGTAAAGAGCACCTGCGCAATGATG      | 1020 |  |  |
| DB              | 1032 | TATTACCTTGGTTTCAAGTCTTTTTCAGACTCATGATAGTAAAGAGCACCTGCGCAATGATG      | 1091 |  |  |
| QY              | 1021 | GAAAGAAATTTAGGACCCATACCAACACACATGATTTAGAAAACAGAAAACGCAAGTAT         | 1080 |  |  |
| DB              | 1092 | GAAAGAAATTTAGGACCCATACCAACACACATGATTTAGAAAACAGAAAACGCAAGTAT         | 1151 |  |  |
| QY              | 1081 | TTTTCACCAATACCAAGTATGATTTGGGATGAACACACAGTCTTCTGCTGGTAGATATGTTAGGAGA | 1140 |  |  |
| DB              | 1152 | TTTTCACCAATACCAAGTATGATTTGGGATGAACACACAGTCTTCTGCTGGTAGATATGTTAGGAGA | 1211 |  |  |
| QY              | 1141 | CGCTGCAAAACCGTTGAAGGAATTTATGCTTTTGTCTCATGATGAAGAACATGAGAAACCTGTTT   | 1200 |  |  |
| DB              | 1212 | CGCTGCAAAACCGTTGAAGGAATTTATGCTTTTGTCTCATGATGAAGAACATGAGAAACCTGTTT   | 1271 |  |  |
| QY              | 1201 | GACCTGGTTCGAAGAACTGTTAGAAATATCATCAACTCAAGAAATTTACCTTGGATCAAGCA      | 1260 |  |  |
| DB              | 1272 | GACCTGGTTCGAAGAACTGTTAGAAATATCATCAACTCAAGAAATTTACCTTGGATCAAGCA      | 1331 |  |  |
| QY              | 1261 | TTGAGACATCCTTTCTTTGACTTATTTAAAAAGAAATGAAATGGGAATCAGTGGTCTTTAC       | 1320 |  |  |
| DB              | 1332 | TTGAGACATCCTTTCTTTGACTTATTTAAAAAGAAATGAAATGGGAATCAGTGGTCTTTAC       | 1391 |  |  |
| QY              | 1321 | TATATACCTCTAGAGAGATTTACTTAAAGCTGTGTCACTCAACATTAACATTTCTAATAT        | 1380 |  |  |
| DB              | 1392 | TATATACCTCTAGAGAGATTTACTTAAAGCTGTGTCACTCAACATTAACATTTCTAATAT        | 1451 |  |  |
| QY              | 1381 | TTTTGTAAACATTAATTTTGTACAGTTAAAGTGAATAATTTGATGTTTGTGATCAAA           | 1440 |  |  |
| DB              | 1452 | TTTTGTAAACATTAATTTTGTACAGTTAAAGTGAATAATTTGATGTTTGTGATCAAA           | 1511 |  |  |
| QY              | 1441 | TAGCATATTAACCTTGTAAAGCAAGTATGGTCTTGATTAATGCAATTAGAAAAATTTAAATTT     | 1500 |  |  |
| DB              | 1512 | TAGCATATTAACCTTGTAAAGCAAGTATGGTCTTGATTAATGCAATTAGAAAAATTTAAATTT     | 1571 |  |  |
| QY              | 1501 | AATTTTCTCTTTGAAATTCATTTTAAATACTTTTAAATACTTTTGAATAATCTTTTGTGTCAGTG   | 1560 |  |  |
| DB              | 1572 | AATTTTCTCTTTGAAATTCATTTTAAATACTTTTAAATACTTTTGAATAATCTTTTGTGTCAGTG   | 1631 |  |  |
| QY              | 1561 | ATAAATGTGATGATCTTGGCTTTTGTACATGGAGGTCCCTCTCAAGTGATTTTGTGTTTG        | 1620 |  |  |
| DB              | 1632 | ATAAATGTGATGATCTTGGCTTTTGTACATGGAGGTCCCTCTCAAGTGATTTTGTGTTTG        | 1691 |  |  |
| QY              | 1621 | AGTAAAGGAAATCTTGACTACTTTTATATCTTTAAAGGAATATTTCTTTATATACCTCAA        | 1680 |  |  |
| DB              | 1692 | AGTAAAGGAAATCTTGACTACTTTTATATCTTTAAAGGAATATTTCTTTATATACCTCAA        | 1751 |  |  |
| QY              | 1681 | TTTAGAACTTAACTTTAAAGTTTTCTTCTGTAATTTGTTGAAACGGGTGATTTATTAA          | 1740 |  |  |
| DB              | 1752 | TTTAGAACTTAACTTTAAAGTTTTCTTCTGTAATTTGTTGAAACGGGTGATTTATTAA          | 1811 |  |  |
| QY              | 1741 | CTCTAGATAAGCAGTACTAGAAACCAAACTCTAGAAAATGTTTACTGTTAGAAATCTAT         | 1800 |  |  |
| DB              | 1812 | CTCTAGATAAGCAGTACTAGAAACCAAACTCTAGAAAATGTTTACTGTTAGAAATCTAT         | 1871 |  |  |
| QY              | 1801 | TAAATTTTAAAGTGTGTTATCTTTTCTTCAATGGGTGATGTCAAGGTGATAACAGACATTC       | 1860 |  |  |
| DB              | 1872 | TAAATTTTAAAGTGTGTTATCTTTTCTTCAATGGGTGATGTCAAGGTGATAACAGACATTC       | 1931 |  |  |
| QY              | 1861 | ATGGAAGGCATGAGTTTGTCCATTTGTGACAGTTTGTGTAATAAACAACACATACACT          | 1920 |  |  |
| DB              | 1932 | ATGGAAGGCATGAGTTTGTCCATTTGTGACAGTTTGTGTAATAAACAACACATACACT          | 1991 |  |  |
| QY              | 1921 | TTATTTAGATTTAAATCTTAATCTGAAAGTCACTTGGAAATGACATTTTCCAGTATG           | 1980 |  |  |
| DB              | 1992 | TTATTTAGATTTAAATCTTAATCTGAAAGTCACTTGGAAATGACATTTTCCAGTATG           | 2051 |  |  |
| QY              | 1981 | TTTGTGTAGTCAAGATATAAATAAGAAATTTCTGATGAGAGTTTTCAGTTTATTAATACC        | 2040 |  |  |
| DB              | 2052 | TTTGTGTAGTCAAGATATAAATAAGAAATTTCTGATGAGAGTTTTCAGTTTATTAATACC        | 2111 |  |  |
| QY              | 2041 | AAGTCTTTAGGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTA      | 2100 |  |  |

```
Db 2112 AAGTCCTTAGGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAA 2171
Qy 2101 ACCATATAGAAATTAAGTTTATTAATTAAGCAATTTATGCTGTGTGATAATTTCTTACGGGAG 2160
Db 2172 ACCATATAGAAATTAAGTTTATTAATTAAGCAATTTATGCTGTGTGATAATTTCTTACGGGAG 2231
Qy 2161 AAAGAGGATTTGATTTGGAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTT 2220
Db 2232 AAAGAGGATTTGATTTGGAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTT 2291
Qy 2221 AATTCATTTGGTTACATAAATTTTTCGACTTCAGAAA 2256
Db 2292 AATTCATTTGGTTACATAAATTTTTCGACTTCAGAAA 2327

RESULT 3
US-10-339-656-1
; Sequence 1, Application US/10339656
; Patent No. 6733978
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-1

Query Match 100.0%; Score 2256; DB 4; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACTCATCCAGTCATATTTAGAACGAGGTCTTGAATGACGCGAGATTTACGGGAC 60
Db 72 GACACTCATCCAGTCATATTTAGAACGAGGTCTTGAATGACGCGAGATTTACGGGAC 131
Qy 61 CGGAGATACGTTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTAT 120
Db 132 CGGAGATACGTTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTAT 191
Qy 121 CACAGAGACATTTGAAGCGGGTATCGAATCCACTCCAGTAAATCTTCAGTCCGACGAGG 180
Db 192 CACAGAGACATTTGAAGCGGGTATCGAATCCACTCCAGTAAATCTTCAGTCCGACGAGG 251
Qy 181 AGAAGCAGTCTTAAAGAGGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAG 240
Db 252 AGAAGCAGTCTTAAAGAGGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAG 311
Qy 241 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGAGGGTCACTGATCTGT 300
Db 312 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGAGGGTCACTGATCTGT 371
Qy 301 CAAAGTGGAGCGTTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGGAGCC 360
Db 372 CAAAGTGGAGCGTTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGGAGCC 431
Qy 361 TTGCGCAAGTTGTAGAGTGCATTCATCATGCGATGGATGGCATGCGATGAGCAGTGAAA 420
Db 432 TTGCGCAAGTTGTAGAGTGCATTCATCATGCGATGGATGGCATGCGATGAGCAGTGAAA 491
```

```
Qy 421 ATCGTAAATAATGTAGCCGTTTACCGTGAACAGAGCTCGTTGAGAAATCAAAGTATTAGAG 480
Db 492 ATCGTAAATAATGTAGCCGTTTACCGTGAACAGAGCTCGTTGAGAAATCAAAGTATTAGAG 551
Qy 481 CACTTAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAGATGGTTT 540
Db 552 CACTTAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAGATGGTTT 611
Qy 541 GATCATCATGTCTATGTTTGTATTGTTTGAACCTACTCTGGGACTTTAGTACTTACGATTTT 600
Db 612 GATCATCATGTCTATGTTTGTATTGTTTGTGTTGAACTACTCTGGGACTTTAGTACTTACGATTTT 671
Qy 601 ATTAAGAAAAACAGCTTTCTCCCATTTCAAATGACCAATCAGCAGATGCGGTATCAG 660
Db 672 ATTAAGAAAAACAGCTTTCTCCCATTTCAAATGACCAATCAGCAGATGCGGTATCAG 731
Qy 661 ATCTGCCAGTCAATAAATTTTTCATCATATAATAAATTAACCCATACAGATCTGAGGCT 720
Db 732 ATCTGCCAGTCAATAAATTTTTCATCATATAATAAATTAACCCATACAGATCTGAGGCT 791
Qy 721 GAAAATATTTTGTGTGAAGTCTGACTATCTAGTCAAAATATAATCTTAAATGAAACGT 780
Db 792 GAAAATATTTTGTGTGAAGTCTGACTATCTAGTCAAAATATAATCTTAAATGAAACGT 851
Qy 781 GATGAACGCACACTGAAACACACAGATATCAAAGTTGTTGACTTTGGAAGTGCACACGTAT 840
Db 852 GATGAACGCACACTGAAACACACAGATATCAAAGTTGTTGACTTTGGAAGTGCACACGTAT 911
Qy 841 GATGATGAACATCAGTACTTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAAT 900
Db 912 GATGATGAACATCAGTACTTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAAT 971
Qy 901 TTGGCTTTTGTGTGTGATGCTTCAGCCTTTGAGTCTTTGGAGCATAGGTTGCAATCTTATTGAA 960
Db 972 TTGGCTTTTGTGTGTGATGCTTCAGCCTTTGAGTCTTTGGAGCATAGGTTGCAATCTTATTGAA 1031
Qy 961 TATTACCTTTGTTTTCACAGTCTTTTTCAGATCATGATAGTAAAGAGCACCTGGCAATGATG 1020
Db 1032 TATTACCTTTGTTTTCACAGTCTTTTTCAGATCATGATAGTAAAGAGCACCTGGCAATGATG 1091
Qy 1021 GAACGAATATTAGGACCATACACACATGATTTCAGAAAAACAGAAAAACGCAAGTAT 1080
Db 1092 GAACGAATATTAGGACCATACACACATGATTTCAGAAAAACAGAAAAACGCAAGTAT 1151
Qy 1081 TTTTACCATAACACGACTAGATTGGATGAACACAGTCTCTGCTGTGATAGTATGTTAGGAGA 1140
Db 1152 TTTTACCATAACACGACTAGATTGGATGAACACAGTCTCTGCTGTGATAGTATGTTAGGAGA 1211
Qy 1141 CGCTGCAAAACCGTTTGAAGGAATTTTATGCTTTTGTCTATGATGAAGAACATGAGAAAACTGTTT 1200
Db 1212 CGCTGCAAAACCGTTTGAAGGAATTTTATGCTTTTGTCTATGATGAAGAACATGAGAAAACTGTTT 1271
Qy 1201 GACCTGGTTTGAAGAAATGTTTGAATATGATCAACTCAAGAAATTTACCTTGGATGAAGCA 1260
Db 1272 GACCTGGTTTGAAGAAATGTTTGAATATGATCAACTCAAGAAATTTACCTTGGATGAAGCA 1331
Qy 1261 TTGCAGCATCTCTTTCTTTGACTTTTAAAAAAGAAATGAAATGGGAATCAGTGGTCTTAC 1320
Db 1332 TTGCAGCATCTCTTTCTTTGACTTTTAAAAAAGAAATGAAATGGGAATCAGTGGTCTTAC 1391
Qy 1321 TATATACCTCTCTAGAGAGATTTACTTAAAGACTGTGTGCTCAGTCAATAAACAATTTCTAATAT 1380
Db 1392 TATATACCTCTCTAGAGAGATTTACTTAAAGACTGTGTGCTCAGTCAATAAACAATTTCTAATAT 1451
Qy 1381 TTTTGTAAACATTAATAATTTTGTAGATTAAGTGTAAATATGTTATGTTTGTATCAAA 1440
Db 1452 TTTTGTAAACATTAATAATTTTGTAGATTAAGTGTAAATATGTTATGTTTGTATCAAA 1511
Qy 1441 TAGCATTAATTAACCTTTGTTTAAAGCAAGTATGTTGCTTGTATATGCATTAGAAAAATTAATAAT 1500
Db 1512 TAGCATTAATTAACCTTTGTTTAAAGCAAGTATGTTGCTTGTATATGCATTAGAAAAATTAATAAT 1571
```

```
QY 1501 AATTTTCTTTTGAATTAACATTTTAAATACTTTGAATAATCTTTGTGTCACGTG 1560
Db 1572 AATTTTCTTTTGAATTAACATTTTAAATACTTTGAATAATCTTTGTGTCACGTG 1631
QY 1561 ATAATGTGATGATCTTGCCTTTTGTACATGGAGTCACCTCTGAAGTGATTTTTT 1620
Db 1632 ATAATGTGATGATCTTGCCTTTTGTACATGGAGTCACCTCTGAAGTGATTTTTT 1691
QY 1621 AGTAAAGGAAATCTTGACTACTTTATATCTTAAAGGAATATCTTTATATACCTCAA 1680
Db 1692 AGTAAAGGAAATCTTGACTACTTTATATCTTAAAGGAATATCTTTATATACCTCAA 1751
QY 1681 TTTAGAACTTAACCTTTAAAGTTTTTCTTCTGTAATTTGTGAACGGGTGATTTATTA 1740
Db 1752 TTTAGAACTTAACCTTTAAAGTTTTTCTTCTGTAATTTGTGAACGGGTGATTTATTA 1811
QY 1741 CTCTAGATAAGCAGTACTAGAACCAAACTCAGAAAATGTTTACTGTTAGAAATTCAT 1800
Db 1812 CTCTAGATAAGCAGTACTAGAACCAAACTCAGAAAATGTTTACTGTTAGAAATTCAT 1871
QY 1801 TAAATTTTAAAGTGTGATTTCTTTTCAATGGGTGATGTCAGGGTGATAACAGACATTC 1860
Db 1872 TAAATTTTAAAGTGTGATTTCTTTTCAATGGGTGATGTCAGGGTGATAACAGACATTC 1931
QY 1861 ATGAAAGGCATGCAAGTTTGCCATTTGACAGTTTGTAAATAAAACACATACACT 1920
Db 1932 ATGAAAGGCATGCAAGTTTGCCATTTGACAGTTTGTAAATAAAACACATACACT 1991
QY 1921 TTATTTAGATTAATCTTAAGTGAAGTCAAGTCAAGTGGAAATGACATTTTCAAGTATG 1980
Db 1992 TTATTTAGATTAATCTTAAGTGAAGTCAAGTCAAGTGGAAATGACATTTTCAAGTATG 2051
QY 1981 TTTGTGAGTCACAGATATAAAATAGAAATTTCTGATGAGAGGTTTCAAGTTTTTAATACC 2040
Db 2052 TTTGTGAGTCACAGATATAAAATAGAAATTTCTGATGAGAGGTTTCAAGTTTTTAATACC 2111
QY 2041 AAGTCCTTAGAGTCTTAACATTTGCGCAGCATCTGTTTATCAATGACATAAATACGTA 2100
Db 2112 AAGTCCTTAGAGTCTTAACATTTGCGCAGCATCTGTTTATCAATGACATAAATACGTA 2171
QY 2101 ACCTATAGAAATTAAGTTTATTAATAGGCAATTTATGTCGTGATTAATCTTACGGGAG 2160
Db 2172 ACCTATAGAAATTAAGTTTATTAATAGGCAATTTATGTCGTGATTAATCTTACGGGAG 2231
QY 2161 AAAGAGGATTTGATTTGAAAGCAGTTTTCGGAAGAAAGTCTGCTGAAATTTTCCAGAAATTT 2220
Db 2232 AAAGAGGATTTGATTTGAAAGCAGTTTTCGGAAGAAAGTCTGCTGAAATTTTCCAGAAATTT 2291
QY 2221 AATTGATTTGTTACATAAACTTTTGTGACTTCAGAAA 2256
Db 2292 AATTGATTTGTTACATAAACTTTTGTGACTTCAGAAA 2327
```

## RESULT 4

```
US-09-016-000-9
; Sequence 9, Application US/09016000
; Patent No. 5962232
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Akerman, Ingrid E.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
```

```
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,000
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0465 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2446 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: NEUTPMT01
CLONE: 339963
US-09-016-000-9
```

Query Match 90.7%; Score 2046.6; DB 2; Length 2446;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 2149; Conservative 0; Mismatches 4; Indels 90; Gaps 1;

```
QY 14 GTCATTTTAAAGCAAGGTCCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 73
Db 288 GTCATTTTAAAGCAAGGTCCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 347
QY 74 ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTATCACAGACATTTG 133
Db 348 ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTATCACAGACATTTG 407
QY 134 AAAGCGGTATCGAATCCACTGCGAGTAATCTTCAGTCCGACGAGGAGAGACGCTCTTA 193
Db 408 AAAGCGGTATCGAATCCACTGCGAGTAATCTTCAGTCCGACGAGGAGAGACGCTCTTA 467
QY 194 AAAGGAAGCGCAATAGACATCTGTTCAAGTCAATCAGTCAAGTTCGAGAGCCACCGAAGGA 253
Db 468 AAAGGAAGCGCAATAGACATCTGTTCAAGTCAATCAGTCAAGTTCGAGAGCCACCGAAGGA 515
QY 254 AAAGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGCG 313
Db 516 ----- 515
QY 314 TTCTAAGAGCAAGATATGAAATCGTGACACCTTTGGGTGAAGGAGCCTTTGGCAAAGTTG 373
Db 516 -----AATCCGTGGACACTTTGGGTGAAGGAGCCTTTGGCAAAGTTG 557
QY 374 TAGAGTGCAATGATCATGGCATGGATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATG 433
Db 558 TAGAGTGCAATGATCATGGCATGGATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATG 617
QY 434 TAGGCGGTTACCGTGAAGCAGCTGTTCCAGAAATCCAAAGTATTAGAGCAGCTTAATAGTA 493
Db 618 TAGGCGGTTACCGTGAAGCAGCTGTTCCAGAAATCCAAAGTATTAGAGCAGCTTAATAGTA 677
QY 494 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAGTGTGTTTCATCATCATGTGTC 553
Db 678 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAGTGTGTTTCATCATCATGTGTC 737
QY 554 ATGTTTGTGTTGTTTGAACCTACTGGGACTTAGTACTTACGATTTTACGATTTTCAATAAGAAACA 613
```







CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 699:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1456 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: HNT2AGT01  
CLONE: 488842  
US-09-023-655-699

Query Match 52.2%; Score 1178.2; DB 4; Length 1456;  
Best Local Similarity 99.7%; Pred. No. 6.7e-299;  
Matches 1191; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 932 TTGGAGCATAGGTTGCAATCTTATTGAATATTACCTTGGTTTACAGTCTTTCAGACTC 991  
DB 262 TGTAGCGCATAGGTTGCAATCTTATTGAATATTACCTTGGTTTTCAGTCTTTCAGACTC 321  
QY 992 ATGATAGTAAAGAGACACCTGGCAATGATGGAACGAATATTAGGACCCATACACAAACACA 1051  
DB 322 ATGATAGTAAAGAGACACCTGGCAATGATGGAACGAATATTAGGACCCATACACAAACACA 381  
QY 1052 TCATTTCAGAAAACAAAGAAAACGCAAGTATTTTCCACATAACCCAGTAGATTGGGATGAAC 1111  
DB 382 TGATTCAGAAAACAAAGAAAACGCAAGTATTTTCCACATAACCCAGTAGATTGGGATGAAC 441  
QY 1112 ACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTATGCTTT 1171  
DB 442 ACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTATGCTTT 501  
QY 1172 GTCATGATGAAGAACATCAGAAACGTTTGACCTGGTTCGAGAGATGTTAGAAATGATGC 1231  
DB 502 GTCATGATGAAGAACATCAGAAACGTTTGACCTGGTTCGAGAGATGTTAGAAATGATGC 561  
QY 1232 CAACCTCAAAGAAATTAACCTTGAATGAAGCATTCAGCATCCTTTCTTTGACCTATTAAAAA 1291  
DB 562 CAACCTCAAAGAAATTAACCTTGAATGAAGCATTCAGCATCCTTTCTTTGACCTATTAAAAA 621  
QY 1292 AGAAATGAATGGGAATCAGTGGTCTTACTATATATCTTCTAGAGAGATTTACTTAAGA 1351  
DB 622 AGAAATGAATGGGAATCAGTGGTCTTACTATATATCTTCTAGAGAGATTTACTTAAGA 681  
QY 1352 CTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGTACAGTT 1411  
DB 682 CTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGTACAGTT 741  
QY 1412 AAGTGAATATTTGATGTTTGTATCAATAGCATAATTAATCTTGAAGCAAGTATGCT 1471  
DB 742 AAGTGAATATTTGATGTTTGTATCAATAGCATAATTAATCTTGAAGCAAGTATGCT 801  
QY 1472 CTTGATAATGCAATTAAGAAAATTAATAATTTTCTTTTGTAAATTA - CCAATTTTAA 1530  
DB 802 CTTGATAATGCAATTAAGAAAATTAATAATTTTCTTTTGTAAATTAATCCATTTTAA 861  
QY 1531 ATACCTTTGAATATTCCTTTGTCAGTGAATAATGATGATTCCTTGCCTTTTGTACA 1590  
DB 862 ATACCTTTGAATATTCCTTTGTCAGTGAATAATGATGATTCCTTGCCTTTTGTACA 921  
QY 1591 TGGAGGTCACTCTGAAGTGAATTTTGTAGTAAAGAAATCTTGTACTCTTATATTT 1650  
DB 922 TGGAGGTCACTCTGAAGTGAATTTTGTAGTAAAGAAATCTTGTACTCTTATTT 981  
QY 1651 CTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAATTTAAAGTTTTCCTTC 1710

Db 982 CTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAATTTAAAGTTTTCCTTC 1041  
QY 1711 TGTAAATTTGCAACGGGTGATTTATTATTAACCTCTAGATAAGCAGGTACTAGAAAACCAAAA 1770  
Db 1042 TGTAAATTTGCAACGGGTGATTTATTATTAACCTCTAGATAAGCAGGTACTAGAAAACCAAAA 1101  
QY 1771 CTCAGAAAATGTTTACTGTTAGAAATCTTATTAAATTTTAAAGTGTGTTTCTTTTCATT 1830  
Db 1102 CTCAGAAAATGTTTACTGTTAGAAATCTTATTAAATTTTAAAGTGTGTTTCTTTTCATT 1161  
QY 1831 GGGTATGTCAGGGTGATAACACAGACATTTCAATGGAAGGCGCATGCTTGTCCATTGTGA 1890  
Db 1162 GGGTATGTCAGGGTGATAACACAGACATTTCAATGGAAGGCGCATGCTTGTCCATTGTGA 1221  
QY 1891 CAGTTTGTTTAATAAACCACATACACACTTTTATTAAAGATTTAAATCTAACTGGAAGT 1950  
Db 1222 CAGTTTGTTTAATAAACCACATACACACTTTTATTAAAGATTTAAATCTAACTGGAAGT 1281  
QY 1951 CAGTTTGGAAAATGGACATTTTCCCAAGTATGTTTGGTGAAGTCAAGATTAATAAATAGAAA 2010  
Db 1282 CAGTTTGGAAAATGGACATTTTCCCAAGTATGTTTGGTGAAGTCAAGATTAATAAATAGAAA 1341  
QY 2011 TTCTGATGAGAGGTTTCAGTTTAAATACCAAGTCTTAAAGTCTTAAACATTTGSCCAGC 2070  
Db 1342 TTCTGATGAGAGGTTTCAGTTTAAATACCAAGTCTTAAAGTCTTAAACATTTGSCCAGC 1401  
QY 2071 ATCTGTTTATCAAAATGACATAAATACGTAACCTTAAAGATTTAAAGTTTATTAAT 2125  
Db 1402 ATCTGTTTATCAAAATGACATAAATACGTAACCTTAAAGATTTAAAGTTTATTAAT 1456

RESULT 6  
US-09-905-999-26  
; Sequence 26, Application US/09905999  
; Patent No. 6797513  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, Axel  
; APPLICANT: NAVLER, Oliver  
; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS  
; FILE REFERENCE: 038602/0431  
; CURRENT APPLICATION NUMBER: US/09/905,999  
; PRIOR FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: 09/127,248  
; PRIOR FILING DATE: 1999-07-31  
; PRIOR APPLICATION NUMBER: PCT/IB97/00946  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 08/877,150  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 60/034,286  
; PRIOR FILING DATE: 1996-12-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 26  
; LENGTH: 1549  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-905-999-26

Query Match 48.6%; Score 1097.4; DB 4; Length 1549;  
Best Local Similarity 90.6%; Pred. No. 1.1e-277;  
Matches 1170; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 14 GTCATTATTTAGAACAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 73  
DB 239 GTCATTATTTAGAACAGATGCTTGAATGAGAGATTTATCGGACCGGAGATACATTG 298  
QY 74 ACGAATACAGGAATGACTTCTGTGAAGGATATGTTCTTAGACATTATCACAGACATTTG 133  
DB 299 ATGAATACAGAAATGACTTCTGCGAAGGATATGTTCCAAAGACATTACCATAGAGACGTTG 358  
QY 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCACTCGGACGAGGAGAGACGTCCTA 193  
DB 359 AAAGCACTTACCGGATCCATTGCGAGTAAATCTTCACTCGGACGAGGAGAGACGTCCTA 418

QY 194 AAAGGAGCGCAATAGACACTGTTCAAGTCATCAGTCAAGTTCGAAGAGCCACCGAAGGA 253  
DB 419 AGAGAGCGTAATAGACCCCTGTGCAAGTCATCAGTCGCATTCGAGAGCCACCGAAGGA 478  
QY 254 AAAGATCCAGAGGATATAGAGGATGATGAGGAGGTCACTCTGATCTGTCAAAAGTGAGACG 313  
DB 479 AAAGATCCAGAGGATATAGAGGATGATGAGGAGGTCACTCTGATCTGTCAAAAGTGAGACG 538  
QY 314 TTCTAAGAGCAAGATATGAATCTGGACACTTTGGGTGAAGGAGCCTTTGGCAAGTTG 373  
DB 539 TTCTAAGAGCAAGATATGAATCTGGACACTTTAGGTGAAGGAGCCTTTGGCAAGTTG 598  
QY 374 TAGAGTGCAATGATCATCGCATGATGCGCATGTAGCAGTGAATAATCGTAAAAAATG 433  
DB 599 TAGAGTGCAATGATCATCGCATGATGCGCATGTAGCAGTGAATAATCGTAAAAAATG 658  
QY 434 TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 493  
DB 659 TAGGAGGTTACCGGAGGAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTGAACAGCA 718  
QY 494 CTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGTTTGCATCATCATGTC 553  
DB 719 CTGACCCCAACAGTGTCTTCCGATGTGTCAGATGCTAGAAATGTTTGCATCATCATGTC 778  
QY 554 ATGTTTGTATTGTTTGAACCTACTGGACCTTAGTACTTACGATTTCAATTAAGAAAAACA 613  
DB 779 ATGTTTGTATTGTTTGAACCTACTGGACCTTAGTACTTACGATTTCAATTAAGAAAAACA 838  
QY 614 GCTTTCTGCCATTTCAAATTCACCAATCAGCAGATGCGCGTATCAGATCTGCCAGTCAA 673  
DB 839 GTTTTCTGCCATTTCAAATTCACCAATCAGCAGATGCGCGTATCAGATCTGCCAGTCAA 898  
QY 674 TAAATTTTACATCATTAATAATTAACCATACAGATCTGAAGCCTGAATAATTTTGT 733  
DB 899 TAAATTTTACATCATTAATAATTAACCATACAGATCTGAAGCCTGAATAATTTTGT 958  
QY 734 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCAAAATGAAACGCTGATGAAACGACAC 793  
DB 959 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCAAAATGAAACGCTGATGAAACGACAC 1018  
QY 794 TGAATAACACAGATATCAAAATGTTGTGACTTTGGAAGTGCACGATGATGATGAAACATC 853  
DB 1019 TGAATAACACAGATATCAAAATGTTGTGACTTTGGAAGTGCACGATGATGATGAAACATC 1078  
QY 854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 913  
DB 1079 ATAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 1138  
QY 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTTTGAATATTACCTTTGGTT 973  
DB 1139 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTTTGAATATTACCTTTGGTT 1198  
QY 974 TCAGAGTCTTTCAGACTCATATAGTAAAGAGCACTCGGCAATGATGAAACGAAATATTAG 1033  
DB 1199 TCAGAGTCTTTCAGAGCCACGATAGTAAAGAGCACTCGGCAATGATGAAACGAAATATTAG 1258  
QY 1034 GACCCATACCAACACATGATTCAGAAACAGAGAAACGCAAGTATTTTCAACCATACC 1093  
DB 1259 GACCCATACCAACATGATTCAGAAACAGAGAAACGCAAGTATTTTCAACCATACC 1318  
QY 1094 AGCTAGATTGGGATGAACACAGTTCTGCTGCTGATATGTTAGGAGAGCGCTGCAACCGGT 1153  
DB 1319 AGCTAGATTGGGATGAACACAGTTCTGCTGCTGATATGTTAGGAGAGCGCTGCAACCGGT 1378  
QY 1154 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACATGTTTGAAGTCTGTTGAA 1213  
DB 1379 TAAAGGAATTTATGCTGTTGTCATGACGAAGAGCATGAGAGTGTGTTGACCTGTTGAA 1438  
QY 1214 GAATGTTAGATATGATCCCACTCAAGAAATTTACCTTCGATGAGCACTTGCAGCATCCTT 1273  
DB 1439 GAATGTTAGATATGATCCCACTCAAGAAATTTACCTTCGATGAGCACTTGCAGCATCCTT 1498

QY 1274 TCCTTGACTTATTAAAAAGAAATGAAATGG 1304  
DB 1499 TCCTTGACTTATTAAAAAGAAATGAGTGGG 1529  
RESULT 7  
US-09-810-671-3  
; Sequence 3, Application US/09810671  
; Patent No. 6455291  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL000758  
; CURRENT APPLICATION NUMBER: US/09/810,671  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 21234  
; TYPE: DNA  
; ORGANISM: Human  
US-09-810-671-3  
Query Match 48.6%; Score 1097; DB 3; Length 21234;  
Best Local Similarity 100.0%; Pred. No. 5, 4e-277;  
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1157 AGGAATTTATGCTTTGTGCATGATGAAGAACATGAGAAACTGTTTGACCTGTTGGAAGAA 1216  
DB 18138 AGGAATTTATGCTTTGTGCATGATGAAGAACATGAGAAACTGTTTGACCTGTTGGAAGAA 18197  
QY 1217 TGTTAGATATGATCCCACTCAAGAAATACCTTTGGATGAAGCATTCAGCATCCTTTCT 1276  
DB 18198 TGTTAGATATGATCCCACTCAAGAAATACCTTTGGATGAAGCATTCAGCATCCTTTCT 18257  
QY 1277 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGTCTTACTATATCTTCTCTAGA 1336  
DB 18258 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGTCTTACTATATCTTCTCTAGA 18317  
QY 1337 AGAGATTACTTTAAGACTGTGTCACTCAATCAATCAATTTTGTAAACATTTAA 1396  
DB 18318 AGAGATTACTTTAAGACTGTGTCACTCAATCAATCAATTTTGTAAACATTTAA 18377  
QY 1397 TTATTTGTACAGTTAAGTGTAAATATTGTATGTTTGTATCAATAGCATTAATTAACCTTG 1456  
DB 18378 TTATTTGTACAGTTAAGTGTAAATATTGTATGTTTGTATCAATAGCATTAATTAACCTTG 18437  
QY 1457 TTAAGCAAGTATGCTTTGTATAATGCAATAGAAAAATTAATAATTTTCTTTTGA 1516  
DB 18438 TTAAGCAAGTATGCTTTGTATAATGCAATAGAAAAATTAATAATTTTCTTTTGA 18497  
QY 1517 ATTAACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGATGAATGATGATGATC 1576  
DB 18498 ATTAACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGATGAATGATGATGATC 18557  
QY 1577 TTGCTTTTGTACATGGAGTCACTCTGAAGTGAATTTTTTTCAGTAAAAAGGAATCTT 1636  
DB 18558 TTGCTTTTGTACATGGAGTCACTCTGAAGTGAATTTTTTTCAGTAAAAAGGAATCTT 18617  
QY 1637 GACTACTTTTATTTCTTTAAAGGAATATTTCTTTATATCTTCAAAATTTAGAACTTAACTTT 1696  
DB 18618 GACTACTTTTATTTCTTTAAAGGAATATTTCTTTATATCTTCAAAATTTAGAACTTAACTTT 18677  
QY 1697 AAAAGTTTTTCTTCTGTAATTTGTAACGGGTGATTTATTAATCTTAGATAAGCAGGT 1756  
DB 18678 AAAAGTTTTTCTTCTGTAATTTGTAACGGGTGATTTATTAATCTTAGATAAGCAGGT 18737  
QY 1757 ACTAGAACCAAACTCAGAAAAATGTTTACCTGTTAGAAATCTTATTAATTTTAACTGTTG 1816  
DB 18738 ACTAGAACCAAACTCAGAAAAATGTTTACCTGTTAGAAATCTTATTAATTTTAACTGTTG 18797

QY 1817 TATTCCTTTTTCATCGGGTGAATGATGAGGTGATAACACAGACATTCATGGAAGGCATGACG 1876  
Db 18798 TATTCCTTTTTCATCGGGTGAATGAGGTGATAACACAGACATTCATGGAAGGCATGACG 18857  
QY 1877 TTTGTCCTATGACAGTTGTTTAAATAAACCACATACACATTTTATTAAGATTAAA 1936  
Db 18856 TTTGTCCTATGACAGTTGTTTAAATAAACCACATACACATTTTATTAAGATTAAA 18917  
QY 1937 TCTAACTGGAAAGTCAGCTTGGAAATGGACATTTTCCAAAGTATGTTTGGTGAGTCACAGA 1996  
Db 18918 TCTAACTGGAAAGTCAGCTTGGAAATGGACATTTTCCAAAGTATGTTTGGTGAGTCACAGA 18977  
QY 1997 TATAAAATAGAAATTCGTATGAGAGGTTTCAGTTTAAATACCAAGTCCTTACGAGTCT 2056  
Db 18978 TATAAAATAGAAATTCGTATGAGAGGTTTCAGTTTAAATACCAAGTCCTTACGAGTCT 19037  
QY 2057 TAACTTGGCCAGCATCTGTTTATCAATGACATATAATACGTAAACCTTATAGAAATTAAG 2116  
Db 19038 TAACTTGGCCAGCATCTGTTTATCAATGACATATAATACGTAAACCTTATAGAAATTAAG 19097  
QY 2117 TTTTAAATAGGCAATTTATGCTGTGATATCTTACGGGAGAAAGAGGATTTGATTG 2176  
Db 19098 TTTTAAATAGGCAATTTATGCTGTGATATCTTACGGGAGAAAGAGGATTTGATTG 19157  
QY 2177 GAAACAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTGTTACAT 2236  
Db 19158 GAAACAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTGTTACAT 19217  
QY 2237 AAACCTTTTGACTTCAG 2253  
Db 19218 AAACCTTTTGACTTCAG 19234

RESULT 8

US-10-109-854-3  
; Sequence 3, Application US/10109854  
; Patent No. 6630337  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEROF  
; FILE REFERENCE: CL000758D1V  
; CURRENT APPLICATION NUMBER: US/10/109,854  
; PRIOR FILING DATE: 2002-04-01  
; PRIOR APPLICATION NUMBER: 60/227,470  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 09/810,671  
; PRIOR FILING DATE: 2001-03-19  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 21234  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-109-854-3

Query Match 48.6%; Score 1097; DB 4; Length 21234;  
Best Local Similarity 100.0%; Pred. No. 5.4e-277;  
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1157 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTGACCTGGTTCGAAGAA 1216  
Db 18138 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTGACCTGGTTCGAAGAA 18197  
QY 1217 TGTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTCAGCATCCTTTCT 1276  
Db 18198 TGTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTCAGCATCCTTTCT 18257  
QY 1277 TTGACTTTATTAAGAAATGAATGGATGATGATGCTTACTATATATCTTCTCTAGA 1336  
Db 18258 TTGACTTTATTAAGAAATGAATGGATGATGATGCTTACTATATATCTTCTCTAGA 18317

QY 1337 AGAGATTACTTAAGACTGTGTGAGTCAACAACTTCTAATATTTTGTAAACATTAAA 1396  
Db 18318 AGAGATTACTTAAGACTGTGTGAGTCAACAACTTCTAATATTTTGTAAACATTAAA 18377  
QY 1397 TTAATTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATTAATTAACCTG 1456  
Db 18378 TTAATTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATTAATTAACCTG 18437  
QY 1457 TTAAGCAAGTATGGTCTTGATAATGCAATAGAAAAATTAATAATTTTCTTTTGTAA 1516  
Db 18438 TTAAGCAAGTATGGTCTTGATAATGCAATAGAAAAATTAATAATTTTCTTTTGTAA 18497  
QY 1517 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCCAGTGATAAATGATGATGATC 1576  
Db 18498 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCCAGTGATAAATGATGATGATC 18557  
QY 1577 TTGCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTGATTAAGGAATCTT 1636  
Db 18558 TTGCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTGATTAAGGAATCTT 18617  
QY 1637 GACTACTTTTATATCTTAAAGCAATATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT 1696  
Db 18618 GACTACTTTTATATCTTAAAGCAATATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT 18677  
QY 1697 AAAAGTTTTTCTTCTGTAATTTGTTGAACGGGTGATTTATTAACCTTAGATAACAGGT 1756  
Db 18678 AAAAGTTTTTCTTCTGTAATTTGTTGAACGGGTGATTTATTAACCTTAGATAACAGGT 18737  
QY 1757 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATTTCTAATTAATTTTAAGTGTG 1816  
Db 18738 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATTTCTAATTAATTTTAAGTGTG 18797  
QY 1817 TATTCCTTTTTCATTTGGGTGATGTCAGGGTGTATTAACACAGACATTCATGGAAGGCATGACG 1876  
Db 18798 TATTCCTTTTTCATTTGGGTGATGTCAGGGTGTATTAACACAGACATTCATGGAAGGCATGACG 18857  
QY 1877 TTTGTCCAATGTCGACAGTTTGTAAATAAACCACATACACATTTTATTAAGATTAAA 1936  
Db 18858 TTTGTCCAATGTCGACAGTTTGTAAATAAACCACATACACATTTTATTAAGATTAAA 18917  
QY 1937 TCTAACTGGAAGTCAGCTTGGAAAAATGGAACATTTCCAAAGTATGTTTGGTGAGTCACAGA 1996  
Db 18918 TCTAACTGGAAGTCAGCTTGGAAAAATGGAACATTTCCAAAGTATGTTTGGTGAGTCACAGA 18977  
QY 1997 TATAAAATAGAAATTCGTATGAGAGGTTTTCAGTTTAAATACCAAGTCCTTACGAGTCT 2056  
Db 18978 TATAAAATAGAAATTCGTATGAGAGGTTTTCAGTTTAAATACCAAGTCCTTACGAGTCT 19037  
QY 2057 TAACTTGGCCAGCATCTGTTTATCAATGACATATAATACGTAAACCTTATAGAAATTAAG 2116  
Db 19038 TAACTTGGCCAGCATCTGTTTATCAATGACATATAATACGTAAACCTTATAGAAATTAAG 19097  
QY 2117 TTTTAAATAGGCAATTTATGCTGTGATATCTTACGGGAGAAAGAGGATTTGATTG 2176  
Db 19098 TTTTAAATAGGCAATTTATGCTGTGATATCTTACGGGAGAAAGAGGATTTGATTG 19157  
QY 2177 GAAACAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTGTTTACAT 2236  
Db 19158 GAAACAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTGTTTACAT 19217  
QY 2237 AAACCTTTTGACTTCAG 2253  
Db 19218 AAACCTTTTGACTTCAG 19234

RESULT 9

US-10-339-656-3  
; Sequence 3, Application US/10339656  
; Patent No. 6733978  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES



QY 567 GTTGAACACTGCGACTTAGTACTAGATTTCAATTAAGAAACAGCTTCTGCCATT 626  
DB 1223 TTTTGAACATTGGGCACTTAGTACTAGATTTCAATTAAGAAACAGCTTCTGCCATT 1282  
QY 627 TCAAAATGACCACTACAGCAGATGCGCTATCAGATCTGCCAGTCAATAAATTTTTCACA 686  
DB 1283 TCGACTGATCATATCAGAAAGATGGCATATCAGATATGCAATCTGTGAATTTTGTGCA 1342  
QY 687 TCATTAATAATTAACCCATACAGATCTGAAGCTGAAATATTTTGTGTGAAGTCTGA 746  
DB 1343 CAGTAATAAGTTGACTCACACAGACTTAAGCTGAAACATCTTATTGTGCACTCTGA 1402  
QY 747 CTATGTAGTCAATATATTTCTAAATGAACGTGATGATGATGATGATGATGATGATGAT 806  
DB 1403 CTACACAGGCGGTATATCCCAAAATTAAGCTGATGATGATGATGATGATGATGATGAT 1462  
QY 807 TATCAAAAGTTGTGACTTTGGAAGTGCACAGCTATGATGATGATGATGATGATGATGAT 866  
DB 1463 TATTAAGTTGTGACTTTGGAAGTGCACAGCTATGATGATGATGATGATGATGATGAT 1522  
QY 867 GTCTACCCGGCACTACAGACTCCGAGGTCAATTTTGGCTTTAGTTGGTCTCAGCCTTG 926  
DB 1523 ATCTACAAGACTATATAGACACCTGAAGTTATTTTGGCTTTAGTTGGTCTCAGCCTTG 1582  
QY 927 TGATGTTTGGACATAGTTGCAATCTTATGAATATTTACCTTTGTTTACAGTCTTTCA 986  
DB 1583 TGATGCTGGAGCTAGGATGCAATCTTATGAATATTTACCTTTGTTTACAGTCTTTCC 1642  
QY 987 GACTCATGATAGTAAAGAGCACTCGCAATGATGAAAGCAATATTAGGACCCATACCA 1046  
DB 1643 AACACAGATAGTAGGAGCATTTAGCAATGATGAAAGCAATATTAGGACCCATACCA 1702  
QY 1047 ACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCACCAATACAGCTAGATTGGGA 1106  
DB 1703 ACATATGATACAGAAAAACAAGAAACGCAAGTATTTTCCACCAATACAGCTAGATTGGGA 1762  
QY 1107 TGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTAT 1166  
DB 1763 TGAACACAGTTCTGCTGGTAGATATGTTTCAAGACGCTGTAACCTCTGAAGGAATTTAT 1822  
QY 1167 GCTTTGTCATGATGAAGAACATGAAAGCTGTTTCAACCTGTTTCAAGAAATTTAGATA 1226  
DB 1823 GCTTTCTCAAGATGTTGAACATGAGCGTCTTTGACCTCATTCAGAAAAATGTTGAGTA 1882  
QY 1227 TGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGAGCATCTTTCTTTTGAATTTAT 1286  
DB 1883 TGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGAGCATCTTTCTTTTGAATTTAT 1942  
QY 1287 AAAAAAGAAATGAATGGGAATCAGTGTCTTACTATATATCTTCTAGAGAGATTTACT 1346  
DB 1943 GAAGAAAA-----GTATATAGATCTGTAATTTGGACGCTCTCTCGAAGAGA-TCTT 1992  
QY 1347 TAAGACTGTGTCAGTC--AACTAAACATCTAATATTTTGTAAACATTTAAATTTTGTG 1404  
DB 1993 ACAGACTGTATCAGTCTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTCT 2052  
QY 1405 TACAGTTAAGTGAATAATTTGATGTTTGTATCAATAGCATAATTAACCTTTGTTAAAGCAA 1464  
DB 2053 TAACATTTTATTTGCGCATTTTATTTTGTGTTGGTAAATTTGTTTCAATTAAGTACATAG 2112  
QY 1465 GTATGGCTGTGATATGATAGAAAAATTAATAATTTTCTTTTGTAAATTTACCAT 1524  
DB 2113 CTAAGGTAATGAACATCTTTTTCAGTAATTTGTAAGTGTATTTTATTCAGAAATAAATTTT 2172  
QY 1525 TTTTAAATACCTTTGAAATAT 1545  
DB 2173 GTGCTTATGAAGTTGATATGT 2193

RESULT 11

US-09-905-999-22

; Sequence 22, Application US/09905999

; Patent No. 6797513

; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, Axel  
; APPLICANT: NAVLER, Oliver  
; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS  
; FILE REFERENCE: 038602/0431  
; CURRENT APPLICATION NUMBER: US/09/905,999  
; CURRENT FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: 09/127,248  
; PRIOR FILING DATE: 1999-07-31  
; PRIOR APPLICATION NUMBER: PCT/IB97/00946  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 08/877,150  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 60/034,286  
; PRIOR FILING DATE: 1996-12-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 22  
; LENGTH: 1538  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; US-09-905-999-22

Query Match 21.3%; Score 481.2; DB 4; Length 1538;

Best Local Similarity 64.6%; Pred. No. 5.5e-116;

Matches 736; Conservative 0; Mismatches 398; Indels 6; Gaps 1;

QY 143 ATCGAATCCACTGCGAGTAAATCTTTCAGTCCGAGCAGGAGGAGCAGTCTTAAAGGAAAGC 202  
DB 329 ATCGAGAGAAACAGCAGTTACCGAGCCAGCGCAGCGGAGGAGGAAACACAGAGGCGGA 388  
QY 203 GCAATAGACACTGTTCAAGTCACTCAGTCCGCTTTCGAGAGCCACCGAAGGAAAGATCCA 262  
DB 389 GGAGAGCGAGCGGACATTCAGCGCTCATCTTTCACAGCAGCGCG-----GAGAGCCA 442  
QY 263 GGAGTATAGAGATGATGAGGAGGCTCAGCTGATCTGTCAAAGTGGAGACGTTCTTAAGAG 322  
DB 443 AGAGTGTAGAGGAGCAGCGCTGAGGCGCACCTCATCTACCACTCGGGGAGCTGGCTACAG 502  
QY 323 CAAGATATGAATCGTGACACTTTGGGTGAAGGAGCCTTTGGCAAAAGTTGTAGAGTGCA 382  
DB 503 AGCGATATGAATTTGTAAAGCACTTTAGAGAGAGGACCTTCGGCGCGAGTTGTGCAGTGTG 562  
QY 383 TTGATCATGGCATGGATGGCATGATGATGAGTGAAGTAAATCGTAAAAAATGTAGGCCGTT 442  
DB 563 TGGACCATCGCAGGCGGACACACGAGTTGCGCTGAAGATCATTAAGAAATGTGGAAGAT 622  
QY 443 ACCGTGAAGCAGCTGTTTCAGAAATCCAAGTATTAGAGCACTTAATATAGTACTGATCCCA 502  
DB 623 ACAAGGAAGCAGCCCGACTAGAAATCAACGCTGCTGGAGAAAATCAATGAGAAAGATCCCTG 682  
QY 503 ATAGTGTCTCCGATGTCTCCAGATGCTAGAAATGGTTTGTATCATCATGTCATGTTTGTGA 562  
DB 683 ACAACAGAACTCTGTGTCCAGATGTTTGTACTGTTGTACTACCATGGCCACATGTGTGA 742  
QY 563 TTGTTTGTGAACACTACTGGGACTTAGTACTTACGATTTTCATTAAGAAAAACAGCTTTTCGC 622  
DB 743 TCTCTTTGAGCTTCTGGGCTTAGCAGCTTCGATTTCTCTCAAGACAACTACCTACCTGC 802  
QY 623 CATTTCAAAATTTGACCAATCAGGCGAGTGGGTATCAGATCTGCCAGTCAATAAATTTT 682  
DB 803 CTTACCCCATCCCAAGTGGCCCATGAGCTTTCAGCTCTGCGAGGCGCGCTCAAGTTTCC 862  
QY 683 TACATCATATAAATTAACCCATACAGATCTGAAGCTGAAAAATATTTTGTGTTGTAAGT 742  
DB 863 TCCATGATACAGTTGACATACGACCTCAAACTCAAAATATTTCTGTTTGTGAAT 922  
QY 743 CTGACTATGTAGTCAAAATATAATTTCTAAATGAAACGCTGATGAAACGACGACCTGAAAAACA 802  
DB 923 CAGACTACGAATCACCTACAACTTAGAGAGAGAGGAGATGAGCGCAGTGTAAAGAGCA 982  
QY 803 CAGATATCAAAAGTTGTGACTTTTGGAGTGAACCTATGATGATGATGATGATGATGATGAT 862

Db 983 CAGCCGTCGGGTGGTGACTTCGGCAGTGGCCACCTTTGACCAGAAACACCATAGCACCA 1042  
QY TGGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGTTGGTCTCAGC 922  
Db 1043 TTGTCTCCACTCGCCATTTACCGAGCCCCCGAGGTCACTCTGGAGTTGGGCTGGTCACAGC 1102  
QY 923 CTTGTGATTTGGAGCATAGTTGCAATTTCTATTGAAATATTACCTGGTTTTCAGGTCT 982  
Db 1103 CATGGATGATGAGCATAGGCTGCATCACTTTTGAGTACTAGTTGGCTTCAACCTCT 1162  
QY 983 TTCAGACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1042  
Db 1163 TCCAGACCCATGACAAACAGAGAGCATCTAGCCATGATGAAAGGATCCTGGGCTCTGTC 1222  
QY 1043 CACACACATGATTCAGAAACAGAAACAGAAACAGAAACAGAAACAGAAACAGAAACAG 1102  
Db 1223 CTTCTCGATGATCAGAAAGACAGAAACAGAAACAGAAACAGAAACAGAAACAGAAAC 1282  
QY 1103 GGGATGAAACACAGTTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1162  
Db 1283 GGGATGAAACACCTCAGCCGCGCTACGTTCTGTGAGAACTGCAAACTCTGCGGGGT 1342  
QY 1163 TTATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1222  
Db 1343 ATCTGACCTCAGAGCAGAGGACCAACCACTCTTCGATCTGATGATGATGATGATGATG 1402  
QY 1223 AATATGATCCAACTCAAGAAATTTACCTTGGATGAGCAATTCAGCATCTTTCTTTGACT 1282  
Db 1403 AGTATGAGCCTGCTAGCGGCTGACCTTAGTGAAGCCCTTCAGCACTCTTCTTCTGCT 1462

RESULT 12

US-09-905-999-24  
; Sequence 24, Application US/09905999  
; Patent No. 679513  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, Axel  
; APPLICANT: NAVIER, Oliver  
; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS  
; FILE REFERENCE: 038602/0431  
; CURRENT APPLICATION NUMBER: US/09/905,999  
; CURRENT FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: 09/127,248  
; PRIOR FILING DATE: 1999-07-31  
; PRIOR APPLICATION NUMBER: PCT/IB97/00946  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 08/877,150  
; PRIOR FILING DATE: 1997-06-17  
; PRIOR APPLICATION NUMBER: US 60/034,286  
; PRIOR FILING DATE: 1996-12-19  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 24  
; LENGTH: 1787  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-905-999-24

Query Match 20.2%; Score 454.8; DB 4; Length 1787;  
Best Local Similarity 65.0%; Pred. No. 5.1e-109;  
Matches 672; Conservative 0; Mismatches 362; Indels 0; Gaps 0;  
QY 246 CCGAAGGAAAGATCCAGAGTATAGAGGATGATGAGGAGGATCACCCTGATCTGTCAAAG 305  
Db 447 CAGTAAGCGCAGCAGCCGGAGTGTGAAGATGACAGAGGCGCCACCTGGTGTGCGGAT 506  
QY 306 TGGAGACGTTCTAAGCAAGATATGAATCTGTGACACTTTGGTGTGAAGAGCCCTTTGG 365  
Db 507 CGCGGATTTGGCTCCAGAGCGCATATGATGATGATGATGATGATGATGATGATGATGAT 566  
QY 366 CAAAGTTGTAGAGTGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 425  
Db 567 CAAAGTTGTAGAGTGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 626

QY 426 AAAAAATAGCGCTTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAGTATTAGACACTT 485  
Db 627 CCGTAATGTGGCAAGTATCGGGAAGCTCGTCTAGAAATTAATGTCTCAAGAAAT 686  
QY 486 AATAGTACTGATCCCAATAGTGTCTCCGATGATGATGATGATGATGATGATGATGATG 545  
Db 687 CAGGAGAAAGACAAAGGAAATAGTTCTTTTGTGCTGATGATGATGATGATGATGATGAT 746  
QY 546 TCATGCTCATGTTTGTATTGTTTGAACCTACTCGGACTTAGTACTTACGATTTCAATTA 605  
Db 747 CCATGGTTCATATGTCATGCGCTTTTGTAGCTCTCTGGCAAGAACCTTTTGTAGTTCT 806  
QY 606 AAAAAACAGCTTTCTGCGCATTTCAAATTTGACCAATCAGGCGAGATGGCGTATCAGAT 665  
Db 807 GGAGAACAACTTCCAGCCTTACCCCTTACCAATGTCGGGCACATGCGCTTACCACTG 866  
QY 666 CCAGTCAATAAATTTTTCATCATATAATAATAATAATAATAATAATAATAATAATAATA 725  
Db 867 TCATGCCCTTAGATTTCTACACGAGAACCCAGCTGACCCACACAGACTTGAAGCCAGAA 926  
QY 726 TATTTTGTGTGAAGTCTGACTATGATCAAAATATAATAATAATAATAATAATAATAATA 785  
Db 927 CATCTTGTGTGAATTTCTGAGTTTGAACCCCTCTCAATGAGCACAAGAGCTGGAGGA 986  
QY 786 ACGCACACTGAAAAACACAGATATCAAAGTTTGTGACTTTTGGAAAGTGCACGATATGA 845  
Db 987 GAAGTCAGTGAAGAACACCCAGCATCCGAGTGGCAGACTTTGGCAGTGCACGTTTGAC 1046  
QY 846 TGAACATCAGATGATTTGGTGTCTACCGGCATCTACAGACTCCCGAGGTCAATTTTGGC 905  
Db 1047 TGAACATCAGACCACTTGTGGCCACCCGCTCACTACCGCCACCTGAGGTGATCCTTGA 1106  
QY 906 TTTAGTTGGTCTCAGCCTTGTGATGTTTGGAGCATAGTTTGCATTTCTTATTGAATAT 965  
Db 1107 GCTGGCTGGGCAAGCCTTGTGATGTTTGGAGTATCGGCTGCATTTCTTTTGTAGTACT 1166  
QY 966 CTTTGGTTTACAGTCTTTCAGACTCATATGATGATGATGATGATGATGATGATGATGATG 1025  
Db 1167 CCGTGGCTTTACACTCTTCCAGACCCATGAAATAGAGAACACTTTGGTTTATGATGAGAA 1226  
QY 1026 AATATTAGGACCCATACACACACATGATTTTCAGAAACAGAAACAGAAACAGAAACAG 1085  
Db 1227 GATTTAGGACCCATCCCATCACAATGATGATGATGATGATGATGATGATGATGATGAT 1286  
QY 1086 CCATAACCCAGCTAGATTGGGATGAACAGTCTCTGCTGTAGATATGTTAGGAGACGCTG 1145  
Db 1287 CAAAGGGGGCTGTGTTGGGATGAGACAGCTCTGATGGCGGTATGTTGAAGGAGAACTG 1346  
QY 1146 CAAACCGTTGAAGGAATTTATGCTTTTGTATGATGATGATGATGATGATGATGATGATG 1205  
Db 1347 CAAACCTCTGAAGAGTTACATGCTGCAGGACTCCCTTGGAGCATGTGCAGCTGTTTGA 1406  
QY 1206 GTTTCAGAGATGTTAGATATGATCCAACTCAAGATTTACCTTGGATGAGCAATTCGA 1265  
Db 1407 GATGAGGAGGATGTTAGAGTTTCGACCTTGCAGCCATCATTTGGCAGAGAGCCCTT 1466  
QY 1266 GCATCCTTTCTTTG 1279  
Db 1467 GCACCCCTTCTTTG 1480

RESULT 13

US-09-016-434-1439  
; Sequence 1439, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,434

FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1439:

SEQUENCE CHARACTERISTICS:

LENGTH: 1762 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g632971

US-09-016-434-1439

Query Match 19.2%; Score 432.4; DB 4; Length 1762;

Best Local Similarity 63.6%; Pred. No. 3.8e-103;

Matches 658; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

```
QY 246 CCGAAGGAAAGATCCAGGAGTATAGAGATGATGAGGAGGATCACTGATCTGCAAG 305
DB 440 CAGTAAAGCGCAGCGGAGTGTGGAAGATGACAAGGAGGATCACTGTTGCCGGAT 499
QY 306 TGGAGACGTTCTAAGCAGCAAGATATGAATCGTGACACTTTGGTGAAGGAGCCTTTGG 365
DB 500 CGCGGATTTGGCTTCCAGAGCGATATGAGATTGTGGGAACTTGGGTGAAGGACCTTTGG 559
QY 366 CAAAGTTGTAGAGTGCATTTGATCATGGCATGGCATGGCATGGCATGGCATGGCATGG 425
DB 560 CAAAGTTGTAGAGTGCATTTGATCATGGCATGGCATGGCATGGCATGGCATGGCATGG 619
QY 426 AAAAAATAGAGCGTTTACCGTGAAGAGAGTGTTCAGAAATCCAAATTCAGATTCAGATTC 485
DB 620 CGCAACGTTGGGCAAGTACCGGAGGCTGCCGGCTAGAAATCAACGTTGCTCAAAAAAT 679
QY 486 AAATAGTACTGATCCCAATAGTGTCTCCGATGTCAGAGTGTAGATGCTGTTGATCA 545
DB 680 CAAAGGAGGAGCAAGAAAGCAAGTCTCTGTGTCTGTGTCTGTGTCTGTGTCTGTGTCT 739
QY 546 TCATGGTCAATGTTTGTATTTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTA 605
DB 740 CCACGTCACATGTCATCGCTTTGAGTCTCTGGCAAGAACACCTTTGAGTCTCTGAA 799
QY 606 AGAAACAGCTTTCTGCAATTTCAATTTGACACATCAGGCGAGATGGGCTATCAGATCTG 665
DB 800 GGAGAATAACTTCCAGCTTTACCCCTACCAATGTCGGGACATGCGGCTACCACTCTG 859
QY 666 CCAGTCAATAATTTTACATCAATAATAATAATAATAATAATAATAATAATAATAATA 725
DB 860 CCACGCTTGTAGATTTCTGATGAGATCAGCTGACCATACAGACTTGTAAACCTTGAA 919
```

```
QY 726 TATTTTGTGTTGAGTCTGACTATGTAGTCAATATATATTTCTAAATGAACGTTGATGA 785
DB 920 CATCTGTTTGTGAATTTCTGAGTTTGAACCTCTTACAAATGAGCAACAAGAGCTGTGAGGA 979
QY 786 ACGCACACTGAAACACAGATATCAAGTTGTTGACCTTTGGAAGTGAACCTGATGATGA 845
DB 980 GAAGTCAGTGAAGAACACAGCATCCGAGTGGCTGACTTTGGCAGTGCCACATTTGACCA 1039
QY 846 TGAACATCACAGTACTTTTGGTGTCTTACCCGGCAGCTTACAGAGCTCCCGAGGTCATTTTGGC 905
DB 1040 TGAGCACACACACCACTTGTGGCCACCTGCTATCTGCGCCCTGAGGTGATCCTTGA 1099
QY 906 TTTAGTGTGTTCTCAGCTTGTGATGTTTGGAGCATAGTGTGCTTCTTATTTGAATATTA 965
DB 1100 GCTGGGCTGGGCACAGCCCTGTGAGCTGTGGAGCATTTGGCTGCAATTTCTTTGAGTACTA 1159
QY 966 CTTTGGTTTTCACAGTCTTTTACAGCTCATGATAGTAAAGAGCACCTGGCAATGATGGAACG 1025
DB 1160 CCGGGCTTTCACACTTCTTCCAGACCCAGCAAAACCGAGAGCACCTGGTGTGATGATGAGAA 1219
QY 1026 AATATTAGGACCCATFACCAACACATGATTTTCAAGAAAACAAGAAACGCAAGTATTTTCA 1085
DB 1220 GATCTAGGGCCCATCCCATCACATGATCCACCGTACCAGGAGCAGAAATATTTCTA 1279
QY 1086 CATAAACACGCTAGATTTGGGATGAACACAGTTCTGCTGTGTAGATATGTTAGAGAGCTG 1145
DB 1280 CAAAGGGGGCTTAGTTTGGGATGAGAACACAGCTCTGACGCGCGGTATGTGNAAGGAGAACTG 1339
QY 1146 CAAACCGTTGAAGGAATTTATGCTTTTGTTCATGATGAAGAACATGAGAACTGTTTGAACCT 1205
DB 1340 CAAACCTCTGAAGAGTTACATGCTCCAAAGCTCCCTGGAGCACGTTGAGCTGTTTGAACCT 1399
QY 1206 GTTTCGAAGAAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTTGGATGAAGCAATTGCA 1265
DB 1400 GATGAGGAGGATGTTAGAAATTTGACCTGCCCGCCAGGATCACACTGGCCGAGGCTTGTCT 1459
QY 1266 GCATCTTTTCTTTG 1279
DB 1460 GCACCTTTTCTTTG 1473
```

#### RESULT 14

US-09-949-016-2648

; Sequence 2648, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2648

; LENGTH: 1763

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-2648

#### Query Match

Best Local Similarity 19.1%; Score 430.8; DB 4; Length 1763;

Matches 657; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

```
QY 246 CCGAAGGAAAGATCCAGGAGTATAGAGATGATGAGGAGGATCACTGATCTGCAAG 305
DB 440 CAGTAAAGCGCAGCGGAGTGTGGAAGATGACAAGGAGGATCACTTGTGTCGGAT 499
```



```
QY 306 TGGAGACGTTCTTAAGAGCAAGATAGAAATCGTGACACTTTGGCTGAAGAGCCTTTGG 365
Db 500 CGGCGATTGGCTCCNAGAGCGATAGATGATGGGGAACCTGGTGAAGCACCCTTTGG 559
QY 366 CAAGTTGTAGAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 425
Db 560 CAAGTTGTAGAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 619
QY 426 AAAAAATCTAGGCGCTTACCGTGAAGAGCTCGTTTCAAGATCCAAATCCAAATCCAA 485
Db 620 CGGCAACCTGGGCAAGTACCGGAGGCTGCGCGGTAGAAATCAACGTGCTCAAAAAAT 679
QY 486 AAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGATGCTTTGATCA 545
Db 680 CAAGAGAGGACAAAGAAACAAAGTCTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGT 739
QY 546 TCATGTCATGTTGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 605
Db 740 CCAGGTCACATGTGATCGCTTTGAGCTCTGCGGCAAGAACACCTTTGAGTTCTCTGAA 799
QY 606 AGAAACAGCTTCTGCGCATTTCAAAATGACACATCAGGAGAGTGGCGTATCAGATCTG 665
Db 800 GGAGATAACTTCCAGCCTTACCCCTACCATGTCCGGCACATGCGCTACCGACTCTG 859
QY 666 CCAGTCAATAAATTTTACATATAAATTAATTAATTAATTAATTAATTAATTAATTAAT 725
Db 860 CAAGGCGCTTGAATCTGAGTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCT 919
QY 726 TATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 785
Db 920 CATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 979
QY 786 ACGCACATCAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845
Db 980 GAAGTCAGTGAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1039
QY 846 TGAACATCAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
Db 1040 TGAGCACACACACACACACACACACACACACACACACACACACACACACACACACAC 1099
QY 906 TTTAGTTGGTCTCAGCCTTGTGATGTTTGGAGCATAGTTGCTTATTTGATATTA 965
Db 1100 GCTGGCTGGGCACAGCCTGTGACGTCTGGAGCATTTGGCTGCAATTTCTTTGAGTACTA 1159
QY 966 CTTGGTTTCAAGTCTTTTCAAGTCTGATGATGATGATGATGATGATGATGATGATGAT 1025
Db 1160 CCGGGCTTCACTCTTTCCAGACCCACGAAACCCGAGAGCACCTGGTGTGATGGAGAA 1219
QY 1026 AATATTAGGACCCATACCAACACATGATTCAGAAAAACGAAACGCAAGTATTTTCA 1085
Db 1220 GATCTAGGGCCATCCCATCACATGATCCACCGTACCGAGAGCAGAAATTTTCTA 1279
QY 1086 CCATTAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1145
Db 1280 CAAGGGGCGCTAGTTTGGATGAGAACAGCTCTGACGCGCGGTATGTGAAGGAGAACTG 1339
QY 1146 CAACCGTTGAGGAAATTTGCTTTGTCATGATGAAGAACATGAGAACTGTTTGACCT 1205
Db 1340 CAACCTCTGAGAGATTACATGCTCCAGACTCCCTGGAGCACGTCGAGCTGTTTGACCT 1399
QY 1206 GTTTCGAAGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1265
Db 1400 GATGAGGAGGATGTTGAAATTTGACCTGCCAGCGCATCACACTGGCGGAGGCGCTGCT 1459
QY 1266 GCATCCTTTCTTTG 1279
Db 1460 GCACCCCTTCTTTG 1473
```

RESULT 15

US-09-949-016-2649

; Sequence 2649, Application US/09949016

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO: 2649
; LENGTH: 1763
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2649
```

Query Match 19.1%; Score 430.8; DB 4; Length 1763;

Best Local Similarity 63.5%; Pred. No. 1e-102;

Matches 657; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

```
QY 246 CCGAAGGAAAGATCCAGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAAAG 305
Db 440 CAGTAGCGGACAGCCGGAGTGTGGAAGATGACAGGAGGTCACTGTTGTCGGAT 499
QY 306 TGGAGACGTTCTTAAGAGCAAGATGAAATCGTGGACACTTTTGGTGAAGAGCCTTTGG 365
Db 500 CGGCGATTGGCTCCAAAGAGCGATATGATGTTGTGGGAAACCTGGTGAAGCACCCTTTGG 559
QY 366 CAAGTTGTAGAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 425
Db 560 CAAGTGTGTGAGTGTGTTGGACCATGCGCAGAGGAAAGTCTCAGGTTGCGCTGAAGATCAT 619
QY 426 AAAAAATCTAGGCGCTTACCGTGAAGAGCTCGTTTCAAGAAATCCAAATTAATAGAGCACTT 485
Db 620 CCGCAACGTGGGCAAGTACCGGAGGCTGCCGGCTAGAAATCAACGTGCTCAAAAAAT 679
QY 486 AAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGGTTGATCA 545
Db 680 CAAGGAGAGGACAAAGAAACAAAGTCTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGT 739
QY 546 TCATGTCATGTTGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 605
Db 740 CCAGGTCACATGTGCAATGCGCTTTGAGCTCTCGGCAAGAACACCTTTGAGTTCTCTGAA 799
QY 606 AGAAACAGCTTCTTCCGATTTCAAAATGACACATCAGGAGAGTGGCGTATCAGATCTG 665
Db 800 GGAGATAACTTCCAGCCTTACCCCTACCATGTCCGGCACATGCGCTACCGACTCTG 859
QY 666 CCAGTCAATAAATTTTACATATAAATTAATTAATTAATTAATTAATTAATTAATTAAT 725
Db 860 CCACGCGCTTGAATTTCTGATGAGAAATCAGCTGACCCATACAGACTTTGAAACACAGAA 919
QY 726 TATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 785
Db 920 CATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 979
QY 786 ACGCACATCAGAAACACAGATATCAAGTGTGTTGACTTTTGGAAAGTCAAGCTATGATGA 845
Db 980 GAAGTCAGTGAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1039
QY 846 TGAACATCAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
Db 1040 TGAGCACACACACACACACACACACACACACACACACACACACACACACACACACAC 1099
QY 906 TTTAGTTGGTCTCAGCCTTGTGATGTTTGGAGCATAGTTGCTTATTTGATATTA 965
Db 1100 GCTGGCTGGGCACAGCCTGTGACGTCTGGAGCATTTGGCTGCAATTTCTTTGAGTACTA 1159
```





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 01:13:38 ; Search time 1192.6 Seconds  
(without alignments)

11198.198 Million cell updates/sec

Title: US-10-801-671-1\_COPY\_72\_2327

Perfect score: 2256

Sequence: 1 gacactcatccagtcatta.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseqn1980s:\*

2: geneseqn1990s:\*

3: geneseqn2000s:\*

4: geneseqn2001as:\*

5: geneseqn2001bs:\*

6: geneseqn2002as:\*

7: geneseqn2002bs:\*

8: geneseqn2003as:\*

9: geneseqn2003bs:\*

10: geneseqn2003cs:\*

11: geneseqn2003ds:\*

12: geneseqn2004as:\*

13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID       | Description        |
|------------|--------|-------------|--------|-------------|--------------------|
| 1          | 2256   | 100.0       | 2354   | 6 AAD32038  | Aad32038 Human kin |
| 2          | 2215.8 | 98.2        | 2497   | 4 AAF89402  | Aaf89402 Human cel |
| 3          | 2048   | 90.8        | 2488   | 4 AAF44675  | Aaf44675 Novel pro |
| 4          | 2048   | 90.8        | 2488   | 12 ADI29373 | Adi29373 Human MAR |
| 5          | 2046.6 | 90.7        | 2446   | 6 AAX89852  | Aax89852 Human pro |
| 6          | 1632   | 72.3        | 1814   | 13 ADP24393 | Adp24393 PRO polyp |
| 7          | 1629.4 | 72.2        | 1792   | 12 ADI26142 | Adi26142 Human CDN |
| 8          | 1300.6 | 57.7        | 1323   | 3 AAC81772  | Aac81772 Human bea |
| 9          | 1286   | 57.0        | 1446   | 12 ADO07809 | Ado07809 Human pol |
| 10         | 1277.4 | 56.6        | 1881   | 12 ADI26140 | Adi26140 Human CDN |
| 11         | 1185   | 52.5        | 1222   | 3 AAC59283  | Aac59283 Human sec |
| 12         | 1178.2 | 52.2        | 1456   | 11 ADI31373 | Adi31373 Human CDN |
| 13         | 1118.8 | 49.6        | 1865   | 12 ADI26144 | Adi26144 Human CDN |
| 14         | 1097.2 | 48.6        | 1446   | 12 ADO07812 | Ado07812 Mouse pol |
| 15         | 1097   | 48.6        | 21234  | 6 AAD32039  | Aad32039 Human kin |
| 16         | 813.2  | 36.0        | 1455   | 12 ADO07810 | Ado07810 Human pol |
| 17         | 811.6  | 36.0        | 1834   | 12 ADH58708 | Adh58708 Human CDC |
| 18         | 811.6  | 36.0        | 1834   | 13 ADRI4000 | Adri4000 Human NF- |
| 19         | 811.6  | 36.0        | 1834   | 13 ADP24718 | Adp24718 PRO polyp |
| 20         | 768.8  | 34.1        | 1452   | 12 ADO07813 | Ado07813 Mouse pol |

|    |       |      |      |             |                    |
|----|-------|------|------|-------------|--------------------|
| 21 | 768.4 | 34.1 | 906  | 12 ADO07814 | Ado07814 Mouse pol |
| 22 | 669.2 | 29.7 | 1785 | 10 ADG74696 | Adg74696 Human kin |
| 23 | 644.4 | 28.6 | 1743 | 12 ADH58715 | Adh58715 Human CDC |
| 24 | 643.2 | 28.5 | 1750 | 6 ABK84044  | Abk84044 Human CDN |
| 25 | 643.2 | 28.5 | 1750 | 10 ADL24750 | Adl24750 Intestina |
| 26 | 640.6 | 28.4 | 2254 | 12 ADE77073 | Ade77073 Human CDN |
| 27 | 640   | 28.4 | 2516 | 3 AAF21734  | Aaf21734 Human bre |
| 28 | 630.6 | 28.0 | 1643 | 10 ADL24751 | Adl24751 Intestina |
| 29 | 499.6 | 22.1 | 1500 | 10 ACA63031 | ACA63031 Human clk |
| 30 | 499.6 | 22.1 | 1973 | 6 ABL64388  | Abi64388 Stomach c |
| 31 | 499.6 | 22.1 | 1973 | 6 ABN95692  | Abn95692 Gene #219 |
| 32 | 499.6 | 22.1 | 1973 | 12 ADQ15173 | Adq15173 Human can |
| 33 | 499.6 | 22.1 | 1973 | 12 ADQ83230 | Adq83230 Human tum |
| 34 | 499.6 | 22.1 | 1973 | 12 ADQ83910 | Adq83910 Human tum |
| 35 | 499.6 | 22.1 | 1973 | 12 ADQ84393 | Adq84393 Human tum |
| 36 | 499.6 | 22.1 | 1973 | 12 ADQ85141 | Adq85141 Human tum |
| 37 | 499.6 | 22.1 | 1973 | 12 ADQ86323 | Adq86323 Human tum |
| 38 | 499.6 | 22.1 | 2111 | 9 AAD57339  | Aad57339 Human kin |
| 39 | 464.4 | 20.6 | 1296 | 4 AAH46906  | Aah46906 CDNA enco |
| 40 | 464.4 | 20.6 | 1842 | 5 AAS87541  | Aas87541 DNA enco  |
| 41 | 461.6 | 20.5 | 1026 | 12 ADO07807 | Ado07807 Human pol |
| 42 | 458.6 | 20.3 | 1885 | 12 ADJ62819 | Adj62819 Human CDN |
| 43 | 452.2 | 20.0 | 1473 | 12 ADO07811 | Ado07811 Mouse pol |
| 44 | 443   | 19.6 | 443  | 3 AAC81771  | Aac81771 Human bea |
| 45 | 441.4 | 19.6 | 2905 | 8 ACC46217  | Acc46217 Human dit |

ALIGNMENTS

RESULT 1

AAD32038

ID AAD32038 standard; cDNA; 2354 BP.

XX AC AAD32038;

XX DT 18-JUN-2002 (first entry)

XX DE Human kinase protein cDNA.

XX KW Human; kinase protein; enzyme; cytostatic; osteopathic; gene expression;

KW colon-moderately differentiated adenocarcinoma; chromosome mapping; Gene;

KW drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;

XX KW bone osteosarcoma; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT 5'UTR 1..32

FT CDS 33..1370

FT 3'UTR 1371..2354

FT product= "Human kinase protein"

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

FT 1371..2354

assays, identifying modulators of kinase activity or treating disorders characterized by absence or unwanted expression of the protein.

Claim 4; Fig 1; 81pp; English.

The invention relates to isolated human kinase proteins and nucleic acids. The nucleic acid and peptide sequences can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate kinase activity in cells and tissues that express the kinase. The nucleic acids are useful as probes or primers, in constructing recombinant vectors, for expressing antigenic portions of the proteins, chromosome mapping, drug screening, testing an individual for a genotype, and for gene therapy in patients containing cells that are aberrant in kinase gene expression. The proteins may be used in drug screening assays, in the identification of compounds that modulate, stimulate or inhibit kinase activity, in pharmacogenomic analysis, in treating disorders characterised by an absence or unwanted expression of the protein (bone osteosarcoma, or colon-moderately differentiated adenocarcinoma), and in generating antibodies specific for the peptides. Such antibodies can be used to detect the protein *in situ*, *in vitro*, or in cell lysate or supernatant, to isolate and purify the proteins from host cells, pharmacogenomic analysis, tissue typing, and in inhibiting protein function. The present sequence is human protein cDNA. Human kinase protein gene is located on chromosome 5

Sequence 2354 BP; 783 A; 370 C; 472 G; 729 T; 0 U; 0 Other;

Query Match 100.0%; Score 2256; DB 6; Length 2354;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACACTCATCCAGTCATTATTAGAACGACAGGTCCTTGAATGACGAGATTATCGGGAC 60  
72 GACACTCATCCAGTCATTATTAGAACGAGGTCCTTGAATGACGAGATTATCGGGAC 131  
61 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT 120  
132 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT 191  
121 CACAGAGACATTGAAGCGGGTATCGAATCCACTGCAGTAATCTTCAGTCGGCAGCAGG 180  
192 CACAGAGACATTGAAGCGGGTATCGAATCCACTGCAGTAATCTTCAGTCGGCAGCAGG 251  
181 AGAAGCAGTCCTTAAAGGAAGCGCAATAGACTGTTCAGAGTCATCAGTCAGTTCGGAAG 240  
252 AGAAGCAGTCCTTAAAGGAAGCGCAATAGACTGTTCAGAGTCATCAGTCAGTTCGGAAG 311  
241 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT 300  
312 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT 371  
301 CAAAGTGAGAGCGTTCTAAGAGCAAGATAGAAATCGTGGACACTTTGGGTGAAGAGCC 360  
372 CAAAGTGAGAGCGTTCTAAGAGCAAGATAGAAATCGTGGACACTTTGGGTGAAGAGCC 431  
361 TTTGGCAAGTGTAGAGTGCAATTCATCATGCAATGAGTGCATGCAATGAGTGAAGTAA 420  
432 TTTGGCAAGTGTAGAGTGCAATTCATCATGCAATGAGTGCATGCAATGAGTGAAGTAA 491  
421 ATCGTAAAAATGTAGGCGGTTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAATATAGAG 480  
492 ATCGTAAAAATGTAGGCGGTTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAATATAGAG 551  
481 CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGCCATGTCAGATGAGTGGTTT 540  
552 CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGCCATGTCAGATGAGTGGTTT 611  
541 GATCATATGGTCATGTTGTATGTTGTTGAATCTTGAATCTTGAATCTTGAATCTTGAAT 600  
612 GATCATATGGTCATGTTGTATGTTGTTGAATCTTGAATCTTGAATCTTGAATCTTGAAT 671  
601 ATTAAGAAAAACAGCTTTTGCCATTTCAGAAATGACCAATCAGGAGATGGCGTATCAG 660

672 ATTAAGAAAAACAGCTTTTGCCATTTCAAATTGACCACATCAGCGAGATGCGGTATCAG 731  
661 ATCTGCCAGTCAATAAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT 720  
732 ATCTGCCAGTCAATAAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT 791  
721 GAAAAATATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAATGAACGT 780  
792 GAAAAATATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAATGAACGT 851  
781 GATGAACGACACTGAAAAACACAGATATCAAAGTTGTGTGACTTTGGAGTGAACGAT 840  
852 GATGAACGACACTGAAAAACACAGATATCAAAGTTGTGTGACTTTGGAGTGAACGAT 911  
841 GATGATGAACATCAGACTCTTTCAGTCTTACCCGCGACTACAGACTCCCGAGTCAAT 900  
912 GATGATGAACATCAGACTCTTTCAGTCTTACCCGCGACTACAGACTCCCGAGTCAAT 971  
901 TTGGCTTTAGGTTGGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTATTGAA 960  
972 TTGGCTTTAGGTTGGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTATTGAA 1031  
961 TATTACCTTTGGTTTACAGTCTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATG 1020  
1032 TATTACCTTTGGTTTACAGTCTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATG 1091  
1021 GAAACGAAATATTAGGACCCATACACACATGATTCAGAAAAACAAAGAAAAACGCAAGTAT 1080  
1092 GAAACGAAATATTAGGACCCATACACACATGATTCAGAAAAACAAAGAAAAACGCAAGTAT 1151  
1081 TTTTCAACCATACACAGTAGATTGGGATGAAACACAGTTTCTGCTGGTAGATATGTTAGGAGA 1140  
1152 TTTTCAACCATACACAGTAGATTGGGATGAAACACAGTTTCTGCTGGTAGATATGTTAGGAGA 1211  
1141 CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACTGTTT 1200  
1212 CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACTGTTT 1271  
1201 GACCTGGTTCGAAGAAATGTTAGAAATATGATCAACTCAAGAAATTTACCTTCGGTGAAGCA 1260  
1272 GACCTGGTTCGAAGAAATGTTAGAAATATGATCAACTCAAGAAATTTACCTTCGGTGAAGCA 1331  
1261 TTGAGCATCCTCTTTTGACTTATTAATAAAGAAATGAAATGGGAATCAGTGGTCTTAC 1320  
1332 TTGAGCATCCTCTTTTGACTTATTAATAAAGAAATGAAATGGGAATCAGTGGTCTTAC 1391  
1321 TATATACCTTCTAGAGAGATTTACTTAAGACTGTGTGAGTCAACTAAACATTTCTAATAT 1380  
1392 TATATACCTTCTAGAGAGATTTACTTAAGACTGTGTGAGTCAACTAAACATTTCTAATAT 1451  
1381 TTTTGTAAACATTAATTTATTTTGTACAGTTAAAGTAAATATGTTATGTTTGTATCAA 1440  
1452 TTTTGTAAACATTAATTTATTTTGTACAGTTAAAGTAAATATGTTATGTTTGTATCAA 1511  
1441 TAGCATATTAACCTTTGTTAAGCAAGTATGGTCTTGAATATGCAATPAGAAAAATTAAT 1500  
1512 TAGCATATTAACCTTTGTTAAGCAAGTATGGTCTTGAATATGCAATPAGAAAAATTAAT 1571  
1501 AATTTTCTTTTGAATATACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTG 1560  
1572 AATTTTCTTTTGAATATACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTG 1631  
1561 ATAAATGTATGTATCTTTGCTTTGTATCATGGAGTCACTCTGAAAGTATTTTTTTTGTG 1620  
1632 ATAAATGTATGTATCTTTGCTTTGTATCATGGAGTCACTCTGAAAGTATTTTTTTTGTG 1691  
1621 AGTAAAGGAAATCTTGACTACTTTATATTTCTTAAAGAAATATTTCTTTATATATCTTCAA 1680  
1692 AGTAAAGGAAATCTTGACTACTTTATATTTCTTAAAGAAATATTTCTTTATATATCTTCAA 1751  
1681 TTTAGAACTTAACCTTTAAAGAGTCTTCTGTAATCTTGAACGGGATGATTATTATTA 1740



Db 1111 ACAGTACTTTGGTGTCTACCGGCACATACAGAGCTCCCGAGGTCATTTTGGCTTAGGTT 1170  
Qy 914 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTGTCATTTCTTATTTGAATATTACCTTGGTT 973  
Db 1171 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTGTCATTTCTTATTTGAATATTACCTTGGTT 1230  
Qy 974 TCACAGTCTTTTCAGACTCATGATAGTAAGACGACCTCGCAATGATGGAACGAATATTAG 1033  
Db 1231 TCACAGTCTTTTCAGACTCATGATAGTAAGACGACCTCGCAATGATGGAACGAATATTAG 1290  
Qy 1034 GACCCATACCAACACACATGATTTTCAGAAAAACAAGAAAAACGCAAGTATTTCACCATTAACC 1093  
Db 1291 GACCCATACCAACACACATGATTTTCAGAAAAACAAGAAAAACGCAAGTATTTCACCATTAACC 1350  
Qy 1094 AGCTAGATTGGATGAACACAGTCTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1153  
Db 1351 AGCTAGATTGGATGAACACAGTCTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1410  
Qy 1154 TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAACTGTTGACCTGGTTGCAA 1213  
Db 1411 TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAACTGTTGACCTGGTTGCAA 1470  
Qy 1214 GAATGTTAGAAATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGCAATTCGAGCATCCTT 1273  
Db 1471 GAATGTTAGAAATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGCAATTCGAGCATCCTT 1530  
Qy 1274 TCTTTGACTTATTAATAAGAAATGAATGGGAATCAGTGGTCTTACTATATCTTCTCT 1333  
Db 1531 TCTTTGGCTTATTAATAAGAAATGAATGGGAATCAGTGGTCTTACTATATCTTCTCT 1590  
Qy 1334 AGAAGAGATTACTTAAGACTGTGTCAGTCACTAAACATTTCTATATTTTGTAAACATT 1393  
Db 1591 AGAAGAGATTACTTAAGACTGTGTCAGTCACTAAACATTTCTATATTTTGTAAACATT 1650  
Qy 1394 AAATTTATTTGTACAGTTAAAGTGAATATTTGTATTTTGTATCAATAGCATATTAAC 1453  
Db 1651 AAATTTATTTGTACAGTTAAAGTGAATATTTGTATTTTGTATCAATAGCATATTAAC 1710  
Qy 1454 TTGTTAAGCAAGTAGTGGTCTGATGAATGATAGAAAAATTAATTAATTTTCTTTT 1513  
Db 1711 TTGTTAAGCAAGTAGTGGTCTGATGAATGATAGAAAAATTAATTAATTTTCTTTT 1770  
Qy 1514 GAAATACCATTTTAAATACCTTTGAAATATCCTTTGTGTCACGATGAATGATGATTG 1573  
Db 1771 GAAATACCATTTTAAATACCTTTGAAATATCCTTTGTGTCACGATGAATGATGATTG 1830  
Qy 1574 ATCTTGCCCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTGATAAAGGAAT 1633  
Db 1831 ATCTTGCCCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTGATAAAGGAAT 1890  
Qy 1634 CTGTACTATTATATCTTAAAGAAATATCTTTATATATCTTCAAATTTAGAACTTAAC 1693  
Db 1891 CTGTACTATTATATCTTAAAGAAATATCTTTATATATCTTCAAATTTAGAACTTAAC 1950  
Qy 1694 TTTAAAGTTTTTCTCTGTAATCTGCAAGGGGATTTATTTAACTCTAGATAAGCA 1753  
Db 1951 TTTAAAGTTTTTCTCTGTAATCTGCAAGGGGATTTATTTAACTCTAGATAAGCA 2010  
Qy 1754 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTCTGTAGAATTTCTATTAATTTTAAAGTG 1813  
Db 2011 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTCTGTAGAATTTCTATTAATTTTAAAGTG 2070  
Qy 1814 TTGTATCTTTTTCATTTGGTGATGTCAGGGTGATTAACGAGCATTCATGGAAGGCATG 1873  
Db 2071 TTGTATCTTTTTCATTTGGTGATGTCAGGGTGATTAACGAGCATTCATGGAAGGCATG 2130  
Qy 1874 CAGTTTGTCCATTGACAGTTTGTGTTAATAAACCACATACACACTTTTATTAAGATTA 1933  
Db 2131 CAGTTTGTCCATTGACAGTTTGTGTTAATAAACCACATACACACTTTTATTAAGATTA 2190  
Qy 1934 AAATCTAACTGGAAGTACGCTTGGAAAAATGGCAATTTCCAGATGATGTTTGGTGAAGTCAC 1993

Db 2191 AAATCTAACTGGAAGTACGCTTGGAAAAATGGCAATTTCCAGATGATGTTTGGTGAAGTCAC 2250  
Qy 1994 AGATATAAAATAGAAAATCTCTGATGAGAGGTTTTCAGTTTTTAAATACCAAGTCTTTAGGAG 2053  
Db 2251 AGATATAAAATAGAAAATCTCTGATGAGAGGTTTTCAGTTTTTAAATACCAAGTCTTTAGGAG 2310  
Qy 2054 TCTTAAACATTTGGCCAGCATCTGTTTATCAATATGATCAATAATACCTTAAGAAATT 2113  
Db 2311 TCTTAAACATTTGGCCAGCATCTGTTTATCAATATGATCAATAATACCTTAAGAAATT 2370  
Qy 2114 AAGTTTATTAATTAGGCAATTTATGCTCTGATATATTTCTACGGGAGAAAGAGATTGA 2173  
Db 2371 AAGTTTATTAATTAGGCAATTTATGCTCTGATATATTTCTACGGGAGAAAGAGATTGA 2430  
Qy 2174 TTGAAAGCAGTTTGGGAAGAACTGCTGCTGAAATTTCCAGAATTTAAATGATTGGTTA 2233  
Db 2431 TTGAAAGCAGTTTGGGAAGAACTGCTGCTGAAATTTCCAGAATTTAAATGATTGGTTA 2490  
Qy 2234 CATAAAC 2240  
Db 2491 CATAAAC 2497  
RESULT 3  
AAF44675  
ID AAF44675 standard; cDNA; 2488 BP.  
XX AAF44675;  
XX  
XX 27-MAR-2001 (first entry)  
XX Novel protein kinase cDNA, SEQ ID NO: 55.  
XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
XX immunosuppressive; cardiac; renal; antiinflammatory; antisthmatic;  
XX dermatological; antidiabetic; antifertility; gene therapy; vaccine;  
XX immune disorder; cardiovascular disease; neurodegenerative disease;  
XX cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
XX inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.  
XX Homo sapiens.  
XX WO200073469-A2.  
XX 07-DEC-2000.  
XX 26-MAY-2000; 2000WO-US014842.  
XX 28-MAY-1999; 99US-0136503P.  
XX (SUGEN-) SUGEN INC.  
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;  
XX WPI; 2001-032161/04.  
XX P-PSDB; AAB65648.  
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
XX treating immune-related diseases and disorders, cardiovascular disease,  
XX neurodegenerative diseases and/or cancers.  
XX Example 1; Fig 2; 310pp; English.  
XX The present sequence encodes a novel protein kinase. The nucleic acids  
XX and the protein kinases they encode may be used in the treatment and  
XX diagnosis of diseases associated with inappropriate kinase expression  
XX such as immune-related diseases and disorders, cardiovascular disease,  
XX neurodegenerative diseases and/or cancers. The nucleic acids and  
XX complementary sequences may also be used as DNA probes in diagnostic  
XX assays. The kinase polypeptides may be used as antigens in the production  
XX of antibodies of kinase expression and activity. Anti-kinase antibodies  
XX and kinase antagonists may also be used to down regulate kinase  
XX expression and activity. Diseases related to kinase expression and

CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure, oxidative-  
CC stress related disorders, chronic inflammatory bowel disease, chronic  
CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,  
CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive  
CC disorders  
XX  
SQ Sequence 2488 BP; 819 A; 400 C; 504 G; 764 T; 0 U; 1 Other;

Query Match 90.8%; Score 2048; DB 4; Length 2488;  
Best Local Similarity 95.9%; Pred No. 0;  
Matches 2148; Conservative 0; Mismatches 1; Indels 90; Gaps 1;

QY 14 GTCATTATTAGAACCAAGGCTCTTGAATGACGAGATATTGCGGACCGGAGATACGTTG 73  
DB 338 GTCATTATTAGAACCAAGGCTCTTGAATGACGAGATATTGCGGACCGGAGATACGTTG 397

QY 74 ACAGATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTTATCACAGACATTTG 133  
DB 398 ACAGATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTTATCACAGACATTTG 457

QY 134 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAGAGACGTCCTTA 193  
DB 458 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAGAGACGTCCTTA 517

QY 194 AAAGGAACGCAATAGACATGTTTCAAGTCATCAGTCACGTTTGAAGAGACCCAGGAGGA 253  
DB 518 AAAGGAACGCAATAGACATGTTTCAAGTCATCAGTCACGTTG----- 561

QY 254 AAAGATCCAGGAGTATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGACG 313  
DB 562 ----- 561

QY 314 TTCTAAGAGCAAGATATGAATCGTGGACACTTTTGGGTGAAGGAGCCTTTGCGAAAGTTG 373  
DB 562 -----NATGAATCGTGGACACTTTTGGGTGAAGGAGCCTTTGCGAAAGTTG 607

QY 374 TAGAGTGCATTGATCATGGCATGGATGCGATGATGATGATGATGATGATGATGATGATGATG 433  
DB 608 TAGAGTGCATTGATCATGGCATGGATGCGATGATGATGATGATGATGATGATGATGATGATG 667

QY 434 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATATAGAGCATTAAATAGTA 493  
DB 668 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATATAGAGCATTAAATAGTA 727

QY 494 CTGATCCCAATAGTCTCTCCGATGTCGAGATGCTAGATGCTAGATGCTAGATGCTAGATGCT 553  
DB 728 CTGATCCCAATAGTCTCTCCGATGTCGAGATGCTAGATGCTAGATGCTAGATGCTAGATGCT 787

QY 554 ATGTTTGTATTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 613  
DB 788 ATGTTTGTATTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 847

QY 614 GCTTTCGCCATTTCAAATTTGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 673  
DB 848 GCTTTCGCCATTTCAAATTTGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 907

QY 674 TAAATTTTATCATCATATAATTAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 733  
DB 908 TAAATTTTATCATCATATAATTAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 967

QY 734 TTGTGAAGTCTGACTATGATGATCAATATTAATTTCAAAATGAAACGTCATGAACGACAC 793  
DB 968 TTGTGAAGTCTGACTATGATGATCAATATTAATTTCAAAATGAAACGTCATGAACGACAC 1027

QY 794 TGAATAACAGATATCAAAAGTTGTTGACTTTTGAAGTGCACCGTATGATGATGATGATGATG 853  
DB 1028 TGAATAACAGATATCAAAAGTTGTTGACTTTTGAAGTGCACCGTATGATGATGATGATGATG 1087

QY 854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCATTTTGGCTTTAGGTT 913  
DB 1088 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCATTTTGGCTTTAGGTT 1147

QY 914 GGTCTCAGCCTTGTGATGTTGGAGCATAGGTGCAATCTTTATTGAATATTACCTTGGTT 973  
DB 1148 GGTCTCAGCCTTGTGATGTTGGAGCATAGGTGCAATCTTTATTGAATATTACCTTGGTT 1207

QY 974 TCACAGTCTTTTCAGACTCATGATAGTAAAGACCACTGGCATGATGGAACGAATATTAG 1033  
DB 1208 TCACAGTCTTTTCAGACTCATGATAGTAAAGACCACTGGCATGATGGAACGAATATTAG 1267

QY 1034 GACCCATACCAACAACACATGATTTTCAAGAAAACGAGAAAACGCAAGTATTTTCCACCAATACC 1093  
DB 1268 GACCCATACCAACAACACATGATTTTCAAGAAAACGAGAAAACGCAAGTATTTTCCACCAATACC 1327

QY 1094 AGCTAGATTGGGATGAACAACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGCACAAACCGT 1153  
DB 1328 AGCTAGATTGGGATGAACAACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGCACAAACCGT 1387

QY 1154 TGAAGGATTTTATGCTTTGTCATGATGAAGACATGAGAAAACGTTTTCACCTGGTTCGAA 1213  
DB 1388 TGAAGGATTTTATGCTTTGTCATGATGAAGACATGAGAAAACGTTTTCACCTGGTTCGAA 1447

QY 1214 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCTT 1273  
DB 1448 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCTT 1507

QY 1274 TCTTTGACTTATTAAAAAGAAATGAAATGGGAATCAGTGGTCTTTATATATATCTTCT 1333  
DB 1508 TCTTTGACTTATTAAAAAGAAATGAAATGGGAATCAGTGGTCTTTATATATATCTTCT 1567

QY 1334 AGAAGAGTTACTTAAAGCTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1393  
DB 1568 AGAAGAGTTACTTAAAGCTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1627

QY 1394 AAATTTATTTTGTACAGTTAAAGTAAATATTGTTGTTTGTATCAATAGCATATTAAC 1453  
DB 1628 AAATTTATTTTGTACAGTTAAAGTAAATATTGTTGTTTGTATCAATAGCATATTAAC 1687

QY 1454 TTGTTAAGCAAGTATGCTCTTGAATAATGCAATAGAAAAATTTAAAAATTTTCTTTT 1513  
DB 1688 TTGTTAAGCAAGTATGCTCTTGAATAATGCAATAGAAAAATTTAAAAATTTTCTTTT 1747

QY 1514 GAAATACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATATGATG 1573  
DB 1748 GAAATACCATTTTAAATATCCTTTGAAATATCCTTTGTGTCAGTGAATATGATG 1807

QY 1574 ATCTTGCTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTGAGTAAAGGAAAT 1633  
DB 1808 ATCTTGCTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTGAGTAAAGGAAAT 1867

QY 1634 CTTGACTACTTTTATATTTCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAAC 1693  
DB 1868 CTTGACTACTTTTATATTTCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAAC 1927

QY 1694 TTTAAAGTTTTCTTCTGTAATTTGTTGAACGGGTGATTTATTTAACTCTAGATAAGCA 1753  
DB 1928 TTTAAAGTTTTCTTCTGTAATTTGTTGAACGGGTGATTTATTTAACTCTAGATAAGCA 1987

QY 1754 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTAAATTTTAAAGTG 1813  
DB 1988 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTAAATTTTAAAGTG 2047

QY 1814 TTGATTCTTTTCTTCTTCAATTTGAGGTGATGTCAGGCTGATTAACAGACATTTCTATGAAAGGCATG 1873  
DB 2048 TTGATTCTTTTCTTCTTCAATTTGAGGTGATGTCAGGCTGATTAACAGACATTTCTATGAAAGGCATG 2107

QY 1874 CAGTTTGTCTTCTGTCAGTTTGTGTTAAATAAAACCATACACACTTTTATTTAAGATTA 1933  
DB 2108 CAGTTTGTCTTCTGTCAGTTTGTGTTAAATAAAACCATACACACTTTTATTTAAGATTA 2167

QY 1934 AAATCTTAACGGAAGTCAGCTTCGAAAAATGACATTTTCCCAAGTATGTTTGTGAGTGCAC 1993  
DB 2168 AAATCTTAACGGAAGTCAGCTTCGAAAAATGACATTTTCCCAAGTATGTTTGTGAGTGCAC 2227

|          |   |  |      |  |  |
|----------|---|--|------|--|--|
| Qy       | 1994  | AGATATATAAATAGAAATTCGTGATGAGAGGTTTCAGTTTTTAATACCAAGTCTCTAGGAG    | 2055 |  |  |
| Db       | 2228  | AGATATATAAATAGAAATTCGTGATGAGAGGTTTCAGTTTTTAATACCAAGTCTCTAGGAG    | 2287 |  |  |
| Qy       | 2054  | TCCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTATAAGAATTT | 2113 |  |  |
| Db       | 2288  | TCCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTATAAGAATTT | 2347 |  |  |
| Qy       | 2114  | AAGTTTATTAATATAGGCAATTTATGCTCTGTGATTAATTTCTTACGGGAGAAAGAGATTGGA  | 2173 |  |  |
| Db       | 2348  | AAGTTTATTAATATAGGCAATTTATGCTCTGTGATTAATTTCTTACGGGAGAAAGAGATTGGA  | 2407 |  |  |
| Qy       | 2174  | TTGGAAGACGAGTTTGGGAGAAAGTGCTGCTGAAATTTCCAGAAATTTAAATTGATTGGTTA   | 2233 |  |  |
| Db       | 2408  | TTGGAAGACGAGTTTGGGAGAAAGTGCTGCTGAAATTTCCAGAAATTTAAATTGATTGGTTA   | 2467 |  |  |
| Qy       | 2234  | CATAAACTTTTTTGACTTCA   | 2252 |  |  |
| Db       | 2468  | CATAAACTTTTTTGACTTCA   | 2486 |  |  |
| RESULT 4 |   |  |      |  |  |
| ID       | ADI29373  |  |      |  |  |
| ID       | ADI29373  | standard; cDNA; 2488 BP.   |      |  |  |
| AC       | XX  |  |      |  |  |
| AC       | ADI29373;   |  |      |  |  |
| XX       | 22-APR-2004   | (first entry)  |      |  |  |
| DT       | XX  |  |      |  |  |
| DE       | XX  | Human MARK3-associated cDNA #40.                                 |      |  |  |
| DE       | XX  |  |      |  |  |
| XX       | XX  | Human; ss; antisense gene therapy; MARK3;                        |      |  |  |
| KW       | KW  | MAP/microtubule affinity-regulating kinase 3; cancer;            |      |  |  |
| KW       | KW  | Alzheimer's disease; neurodegenerative disorder;                 |      |  |  |
| KW       | KW  | hyperproliferative disorder; cytostatic.                         |      |  |  |
| XX       | XX  |  |      |  |  |
| OS       | XX  | homo sapiens.  |      |  |  |
| XX       | XX  |  |      |  |  |
| PN       | US2003232771-A1.  |  |      |  |  |
| XX       | XX  |  |      |  |  |
| PD       | 18-DEC-2003.  |  |      |  |  |
| XX       | XX  |  |      |  |  |
| PF       | 17-JUN-2002; 2002US-00174319.   |  |      |  |  |
| XX       | XX  |  |      |  |  |
| PR       | 17-JUN-2002; 2002US-00174319.   |  |      |  |  |
| XX       | XX  | (ISIS-) ISIS PHARM INC.  |      |  |  |
| PA       | XX  |  |      |  |  |
| XX       | XX  |  |      |  |  |
| FI       | Ward DT, Freier SM, Dobie KW;   |  |      |  |  |
| XX       | XX  |  |      |  |  |
| XX       | WPI; 2004-052188/05.  |  |      |  |  |
| DR       | P-PSDB; ADI29255.   |  |      |  |  |
| XX       | XX  |  |      |  |  |
| PT       | New antisense compound targeted to a nucleic acid molecule encoding       |  |      |  |  |
| PT       | microtubule-affinity-regulating kinases (MARK3), useful for modulating    |  |      |  |  |
| PT       | expression of MARK3 or for treating cancer or Alzheimer's disease.        |  |      |  |  |
| XX       | XX  |  |      |  |  |
| PS       | Disclosure; Fig 2; 233pp; English.  |  |      |  |  |
| XX       | XX  |  |      |  |  |
| CC       | The invention relates to a compound comprising a sequence comprising 8-80 |  |      |  |  |
| CC       | base pairs (bp) targeted to a nucleic acid encoding MARK3                 |  |      |  |  |
| CC       | (MAP/microtubule affinity-regulating kinase 3), that specifically         |  |      |  |  |
| CC       | hybridises with the nucleic acid encoding MARK3; and inhibits expression  |  |      |  |  |
| CC       | of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a  |  |      |  |  |
| CC       | composition comprising the compound and a carrier or diluent, inhibiting  |  |      |  |  |
| CC       | the expression of MARK3 in cells or tissues, treating an animal having or |  |      |  |  |
| CC       | suspected of having a disease or condition associated with MARK3 and      |  |      |  |  |
| CC       | screening for an antisense compound. The antisense oligonucleotide is     |  |      |  |  |
| CC       | useful for preparing a composition for treating hyperproliferative        |  |      |  |  |
| CC       | disorder, particularly cancer and neurodegenerative diseases e.g.         |  |      |  |  |
| CC       | Alzheimer's disease. The present sequence is a MARK3 associated cDNA      |  |      |  |  |
| CC       | included in the figures but not mentioned anywhere else in the            |  |      |  |  |
| CC       | specification.  |  |      |  |  |





QY 74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGAGACATTG 133  
Db 348 ACGAATA CAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGAGACATTG 407  
QY 134 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCACTCCGACGAGGAGAGCAGTCTTA 193  
Db 408 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCACTCCGACGAGGAGAGCAGTCTTA 467  
QY 194 AAAGGAGCGCAATAGACACTGTTTCAAGTCATCAGTCAGTTCGGAAGAGCCACCGAAGGA 253  
Db 468 AAAGGAGCGCAATAGACACTGTTTCAAGTCATCAGTCAGTTCGGAAGAGCCACCGAAGGA 515  
QY 254 AAAGATCCAGGAGTATAGAGGATGATGAGAGGGTCACTCTGTCAAAAGTGAGAGCG 313  
Db 516 ----- 515  
QY 314 TTCTAAGACGAAGATATGAATTCGTGGACACTTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 373  
Db 516 -----AATCCGTGGACACTTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 557  
QY 374 TAGAGTGCAATTGATCATCGCATGGATGCGATGATAGCAGTGAATAATCGTAAATAATG 433  
Db 558 TAGAGTGCAATTGATCATCGCATGGATGCGATGATAGCAGTGAATAATCGTAAATAATG 617  
QY 434 TAGGCCGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 493  
Db 618 TAGGCCGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 677  
QY 494 CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGAAATGGTTTTCATCATCATGCTC 553  
Db 678 CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGAAATGGTTTTCATCATCATGCTC 737  
QY 554 ATGTTTGTATTGTTTGAACCTACTCGGACCTTAGTACTTACGATTTCAATTAAGAAAAACA 613  
Db 738 ATGTTTGTATTGTTTGAACCTACTCGGACCTTAGTACTTACGATTTCAATTAAGAAAAACA 797  
QY 614 GCTTTCTGCCATTTCAAAATGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 673  
Db 798 GCTTTCTGCCATTTCAAAATGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 857  
QY 674 TAAATTTTACATCATATAATTAATTAACCCATACAGATCTGAGCCTGAAAATATTTTGT 733  
Db 858 TAAATTTTACATCATATAATTAATTAACCCATACAGATCTGAGCCTGAAAATATTTTGT 917  
QY 734 TTGTGAAGTCTGACTATGTAGTCAAAATATAATCTTAAATGAAACGCTGATGAACCCAC 793  
Db 918 TTGTGAAGTCTGACTATGTAGTCAAAATATAATCTTAAATGAAACGCTGATGAACCCAC 977  
QY 794 TGAATAAACACAGATATCAAAGTTGTGACTTTGGAAAGTGCAACGATGATGATGAACATC 853  
Db 978 TGAATAAACACAGATATCAAAGTTGTGACTTTGGAAAGTGCAACGATGATGATGAACATC 1037  
QY 854 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGGTT 913  
Db 1038 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGGTT 1097  
QY 914 GGTCTCAGCCTTTGATGTTTGGACATAGGTTGCAATCTTATTGAATATTACCTTTGGTT 973  
Db 1098 GGTCTCAGCCTTTGATGTTTGGACATAGGTTGCAATCTTATTGAATATTACCTTTGGTT 1157  
QY 974 TCACAGTCTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGGAACGAATATTAG 1033  
Db 1158 TCACAGTCTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGGAACGAATATTAG 1217  
QY 1034 GACCCATACCAACACACATGATTGAGAAAACGAAGAACGCAAGTATTTTCCACATACCC 1093  
Db 1218 GACCCATACCAACACACATGATTGAGAAAACGAAGAACGCAAGTATTTTCCACATACCC 1277  
QY 1094 AGCTAGATTGGATGAACACACGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAACCGT 1153  
Db 1278 AGCTAGATTGGATGAACACACGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAACCGT 1337  
QY 1154 TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTGACCTGGTTCGAA 1213

Db 1338 TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTACCTGGTTCGAA 1397  
QY 1214 GAATGTTAGGAATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGACATTTGCAGACATCCTT 1273  
Db 1398 GAATGTTAGGAATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGACATTTGCAGACATCCTT 1457  
QY 1274 TCTTTGACTTATTAAAAAAGAAATGAAATGGGAATCAAGTGGTCTTCTATATATCTTCTCT 1333  
Db 1458 TCTTTGACTTATTAAAAAAGAAATGAAATGGGAATCAAGTGGTCTTCTATATATCTTCTCT 1517  
QY 1334 AGAAGAGATTTACTTAAGACTGTGTCAGTCAACTAAACATTTCTTAATATTTTGTAAACATTT 1393  
Db 1518 AGAAGAGATTTACTTAAGACTGTGTCAGTCAACTAAACATTTCTTAATATTTTGTAAACATTT 1577  
QY 1394 AAATTTATTTCTGACAGTTTAAAGTGAATATTTGATGTTTCTGATCAATAGCATATTAATAC 1453  
Db 1578 AAATTTATTTCTGACAGTTTAAAGTGAATATTTGATGTTTCTGATCAATAGCATATTAATAC 1637  
QY 1454 TTGTTAAGCAAGTATGGTCTTTGATAATGCAATTTAGAAAAATTTAAATTTTCTTTT 1513  
Db 1638 TTGTTAAGCAAGTATGGTCTTTGATAATGCAATTTAGAAAAATTTAAATTTTCTTTT 1697  
QY 1514 GAAATTTACCATTTTAAATACCTTTGAAATATCTTTGTCGCCAGTGAATAATGTGATG 1573  
Db 1698 GAAATTTACCATTTTAAATACCTTTGAAATATCTTTGTCGCCAGTGAATAATGTGATG 1757  
QY 1574 ATCTTGCCCTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTAGTAAAAAGGAAT 1633  
Db 1758 ATCTTGCCCTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTAGTAAAAAGGAAT 1817  
QY 1634 CTTCAGACTCTTTATATTCTTAAAGGAATATTCTTTATATATACTTCAAAATTTAGAACTTAAAC 1693  
Db 1818 CTTCAGACTCTTTATATTCTTAAAGGAATATTCTTTATATATACTTCAAAATTTAGAACTTAAAC 1877  
QY 1694 TTTAAAGTTTTTCTCTGTAATTTGTTGAAACGGGTGATTTATTAACCTTAGATAAGCA 1753  
Db 1878 TTTAAAGTTTTTCTCTGTAATTTGTTGAAACGGGTGATTTATTAACCTTAGATAAGCA 1937  
QY 1754 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTTGAATTTCTATTAATTTTAAAGTG 1813  
Db 1938 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTTGAATTTCTATTAATTTTAAAGTG 1997  
QY 1814 TTGATTTCTTTTTCATTTGGGTGATGTCAGGGTGATTAACACAGACATTCATGGAAGGCAATG 1873  
Db 1998 TTGATTTCTTTTTCATTTGGGTGATGTCAGGGTGATTAACACAGACATTCATGGAAGGCAATG 2057  
QY 1874 CAGTTTGTCCATTTGTGACAGTTTCTTAAATAAAACCATACACACATTTATTTAAGATTA 1933  
Db 2058 CAGTTTGTCCATTTGTGACAGTTTCTTAAATAAAACCATACACACATTTATTTAAGATTA 2117  
QY 1934 AAATCTAACTGGAAGTCAGCTTCGAAAAATGGAACATTTCCAAGTATGTTTGGTGAGTCAC 1993  
Db 2118 AAATCTAACTGGAAGTCAGCTTCGAAAAATGGAACATTTCCAAGTATGTTTGGTGAGTCAC 2177  
QY 1994 AGATATAAAAAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTAAATCAACCAAGTCTTTAGGAG 2053  
Db 2178 AGATATAAAAAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTAAATCAACCAAGTCTTTAGGAG 2237  
QY 2054 TCTTAAACATTTGCCAGCATCTGTTTATCAAAATGATCAATAAATACGTAACCTTATAAGAAAT 2113  
Db 2238 TCTTAAACATTTGCCAGCATCTGTTTATCAAAATGATCAATAAATACGTAACCTTATAAGAAAT 2297  
QY 2114 AAGTTTATTAATTAGGCAATTTATGTCGTGATTAATTTCTTACGGGAGAAAGGATTTGA 2173  
Db 2298 AAGTTTATTAATTAGGCAATTTATGTCGTGATTAATTTCTTACGGGAGAAAGGATTTGA 2357  
QY 2174 TTGGAAGACGTTTGGGAGAAAGAGTGTGCTGCAATTTTCCAGAAATTTAATGATTTGGTTA 2233  
Db 2358 TTGGAAGACGTTTGGGAGAAAGAGTGTGCTGCAATTTTCCAGAAATTTAATGATTTGGTTA 2417  
QY 2234 CATAAACTTTTTGACTTCCAGAAA 2256

|          |      |   |      |
|----------|------|---|------|
| Db       | 2418 | CATAAACTTTTGACTTCAAAA   | 2440 |
| RESULT 6 |      |   |      |
| ADP24393 | ID   | ADP24393 standard; cDNA; 1814 BP.   |      |
| XX       | AC   | ADP24393;   |      |
| XX       | DT   | 18-NOV-2004 (first entry)   |      |
| XX       | DE   | PRO polypeptide encoding cDNA SEQ ID NO:1571.                             |      |
| XX       | DE   | ss; gene; PRO; antiinflammatory; antiarthritic; anti-rheumatic;           |      |
| KW       | DE   | immunosuppressive; osteopathic; antidiabetic; dermatological;             |      |
| KW       | DE   | antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;    |      |
| KW       | DE   | gene therapy; immune system.  |      |
| XX       | OS   | Unidentified.   |      |
| XX       | OS   | WO2004041170-A2.  |      |
| XX       | PN   | 21-MAY-2004.  |      |
| XX       | PD   | 30-OCT-2003; 2003WO-US034312.   |      |
| XX       | PF   | 01-NOV-2002; 2002US-0423394P.   |      |
| XX       | PR   | (GENTH ) GENENTECH INC.   |      |
| XX       | PA   | Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;              |      |
| XX       | PI   | Wu TD;  |      |
| XX       | PI   | WPI; 2004-419628/39.  |      |
| XX       | DR   | P-PSDB; ADP24394.   |      |
| XX       | DR   | New PRO polypeptides and polynucleotides, useful for treating e.g.        |      |
| XX       | PT   | erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated    |      |
| XX       | PT   | renal disease, or demyelinating diseases of the central or peripheral     |      |
| XX       | PT   | nervous system.   |      |
| XX       | PS   | Claim 1; SEQ ID NO 1571; 2940pp; English.                                 |      |
| XX       | CC   | The invention relates to a novel isolated nucleic acid and the PRO        |      |
| XX       | CC   | polypeptide encoded by it. A protein of the invention has                 |      |
| XX       | CC   | antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,        |      |
| XX       | CC   | osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,   |      |
| XX       | CC   | antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide   |      |
| XX       | CC   | of the invention may have a use in gene therapy. The PRO polypeptide, its |      |
| XX       | CC   | agonist, antagonist, or antibody that specifically binds to the           |      |
| XX       | CC   | polypeptide is useful for treating an immune related disorder such as     |      |
| XX       | CC   | systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,        |      |
| XX       | CC   | juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an |      |
| XX       | CC   | idiopathic inflammatory myopathy, Sjogren's syndrome, systemic            |      |
| XX       | CC   | vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune        |      |
| XX       | CC   | thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal   |      |
| XX       | CC   | disease, a demyelinating disease of the central or peripheral nervous     |      |
| XX       | CC   | system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, |      |
| XX       | CC   | a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary      |      |
| XX       | CC   | disease, infectious or autoimmune chronic active hepatitis, primary       |      |
| XX       | CC   | biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,       |      |
| XX       | CC   | inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's       |      |
| XX       | CC   | disease, an autoimmune or immune-mediated skin disease, a bullous skin    |      |
| XX       | CC   | disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  |      |
| XX       | CC   | disease, asthma, allergic rhinitis, atopic dermatitis, food               |      |
| XX       | CC   | hypersensitivity, urticaria, an immunologic disease of the lung,          |      |
| XX       | CC   | eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity   |      |
| XX       | CC   | pneumonitis, a transplantation associated disease, graft rejection or     |      |
| XX       | CC   | graft-versus-host disease. The present sequence encodes a PRO protein of  |      |
| XX       | CC   | the invention.  |      |
| XX       | CC   | Sequence 1814 BP; 609 A; 299 C; 385 G; 520 T; 0 U; 1 Other;               |      |

|    |      |   |      |
|----|------|---|------|
| Db | 1181 | GACCCATACCACAACACATGATTTCAGAAAAACAAGAAACCGAATATTTTTCACCATTAACC  | 1240 |
| Qy | 1094 | AGCTAGATTGGGATGAACACAGTGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT | 1153 |
| Dd | 1241 | AGCTAGATTGGGATGAACACAGTGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT | 1300 |
| Qy | 1154 | TGAAGGAATTTATGCTTTGTTCATGATGAAGAACATAGAGAAACTGTGTGACCTGGTTCGAA  | 1213 |
| Dd | 1301 | TGAAGGAATTTATGCTTTGTTCATGATGAAGAACATAGAGAAACTGTGTGACCTGGTTCGAA  | 1360 |
| Qy | 1214 | GAATGTTAGAATAATGATGCCAACTCAAAGAATTACCTTGGATGAAGCAATTGCAGCATCCTT | 1273 |
| Dd | 1361 | GAATGTTAGAATAATGATGCCAACTCAAAGAATTACCTTGGATGAAGCAATTGCAGCATCCTT | 1420 |
| Qy | 1274 | TCCTTGACTTATTAAAAAGAAATGAATGGGAATCAGTGGTCTTACTATATCTTCTCT       | 1333 |
| Dd | 1421 | TCCTTGACTTATTAAAAAGAAATGAATGGGAATCAGTGGTCTTACTATATCTTCTCT       | 1480 |
| Qy | 1334 | AGAAGACATTACTTTAAGACATGTGTCAGTCAACTAAACATTTCTAATATTTTGTGAAACATT | 1393 |
| Dd | 1481 | AGAAGACATTACTTTAAGACATGTGTCAGTCAACTAAACATTTCTAATATTTTGTGAAACATT | 1540 |
| Qy | 1394 | AAATATTTTGTACAGTTAAGTGPAANAATGTATGTTTTGTATCAATAGCATTAATTAAC     | 1453 |
| Dd | 1541 | AAATATTTTGTACAGTTAAGTGPAANAATGTATGTTTTGTATCAATAGCATTAATTAAC     | 1600 |
| Qy | 1454 | TTGTGTAAGCAAGTATGCTTTGATAATGCAATAGAAAAATTTAAATTTAAATTTTCTTTTT   | 1513 |
| Dd | 1601 | TTGTGTAAGCAAGTATGCTTTGATAATGCAATAGAAAAATTTAAATTTAAATTTTCTTTTT   | 1660 |
| Qy | 1514 | GAATTAACCATTTTTTAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATG      | 1573 |
| Dd | 1661 | GAATTAACCATTTTTTAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATG      | 1720 |
| Qy | 1574 | ATCTTGCTCTTTGTACATGGAGGTCACCTCTGGAAGTCATTTTTTTTTTCAGTAAAGGAAAT  | 1633 |
| Dd | 1721 | ATCTTGCTCTTTGTACATGGAGGTCACCTCTGGAAGTCATTTTTTTTTTCAGTAAAGGAAAT  | 1780 |
| Qy | 1634 | CTTGACTACTTTATATTTCTTTAAAGGAATA                                 | 1662 |
| Dd | 1781 | CTTGACTACTTTTAAAAAAAAAAAAAAAAAAAA                               | 1809 |

## RESULT 7

AD126142  
ID AD126142 standard: cpDNA: 1792 bp.

AD I 26142:

DT 22-APR-2004 (first entry)

XX DE Human cDNA encoding protein that promotes STAT6 activation #54.

XX human CD44 encoding protein and the p150<sup>cas</sup> protein with  
XX  
XX  
KW ST6; immunogen; ST6b activation; allergy; inflammation;  
KW ST6b; immunogen; ST6b activation; allergy; inflammation;  
KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;  
KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;  
KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;  
KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS;  
KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS;  
KW

XX  
OS Homo sapiens.

XX  
PN WO2003104277-A2.

XX  
PD 18-DEC-2003.

05-JUN-2003: 2003WO-JP007123.

05-JUN-2002: 2002JP-00164257.

PR 06-JUN-2002; 2002US-0385912P.  
PR 26-DEC-2002; 2002JP-00377326.

PR 27-DEC-2002; 2002US-0436467P.

|    |  |
|----|--|
| PR | 15-MAY-2003; 2003JP-00137505.                                      |
| PR | 16-MAY-2003; 2003US-0470836P.                                      |
| XX |  |
| XX | (ASAH ) ASAH KASEI KK.   |
| XX |  |
| PI | Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;           |
| XX |  |
| XX | WPI; 2004-122214/12.   |
| DR | P-FSDB; ADI26143.  |
| XX |  |
| PT | New signal transducer and activator of transcription 6 activation  |
| PT | promoting purified protein, for diagnosing and treating disease    |
| PT | associated with activation/inhibition of transcription factor e.g. |
| PT | diabetes and cancer.   |
| PT |  |

PS Claim 4: SEO ID NO 107: 1368pp; English.

CC The invention relates to a purified protein promoting signal transducer  
CC and activator of transcription 6 activation (STAT6). The protein is  
CC useful for the producing an antibody, which involves administering the  
CC protein or its epitope-bearing fragments to a non-human animal as an  
CC antigen. The nucleic acid is useful for diagnosing a disease or  
CC susceptibility to a disease related to expression or activity of the  
CC protein. A transformant expressing the protein is useful for screening  
CC compounds which inhibit or promote STAT6 activation. A transformant  
CC expressing the protein is useful for producing a pharmaceutical  
CC composition. Compositions, antibodies and antisense molecules are useful  
CC for the treating a disease associated with STAT6 activation such as  
CC allergic diseases, inflammation, autoimmune diseases, diabetes,  
CC hyperlipidaemia, infectious disease and cancers. Compositions are useful  
CC for treating disease associated with STAT6 activation and/or prevention  
CC of Th hyperactive diseases. Compositions are also useful in rheumatoid  
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,  
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,  
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6  
CC activity. The protein or nucleic acid is effectively useful for screening  
CC compounds for treating and preventing disease associated with excessive  
CC activation or inhibition of STAT6. The present sequence represents a  
CC human cDNA encoding a protein which promotes STAT6 activation.

Sequence 1792 BP: 587 A; 299 C; 385 G; 520 T; 0 U; 1 Other; XX

Query Match 72.2%; Score 1629.4; DB 12; Length 1792;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 GTCATTATTTAGAACGCAAGGTCCTTGGATGAGCGAGATTATCGGACCGGAGATACGTTG 73

Db  
161 GTCATTATTAGAAGCAAGGTCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG 220

74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGAGACATTG 133

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

103

134 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGAGCAGGAGAAAGCAGTCCIA 193

Db 281 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAAAGCAGTCTTA 340

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

343 333 323 313 303 293 283 273 263 253 243 233 223 213 203 193 183 173 163 153 143 133 123 113 103 93 83 73 63 53 43 33 23 13 3 0

DB 341 AAAGGAAGCGCAATAGACACTGTCAAGTCATCAGTCACGTCGAAAGAGCCACCCGACCGA 400

QY  
254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACCTGATCTGTCTCAAAGTGGAGACG 313

460

TACCCGCGCCTATCATGGTTCCTGTCTCAGAGGAGACG

401 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACCTGATCTGTCATTAATGGAGACG 400

QY  
314 TTCCTAAGAGCAAGATA TGAAATCGTGGACACTTTGGGTGAAGAGCCCTTTGGCAAAGTTG 373

```
QY 434 TAGGCCGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 493
Db 581 TAGGCCGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 640
QY 494 CTGATCCCAATAGTGTCTTCGAGATGTCGAGATGCTAGAAATGGTTGATCATCATGTC 553
Db 641 CTGATCCCAATAGTGTCTTCGAGATGTCGAGATGCTAGAAATGGTTGATCATCATGTC 700
QY 554 ATGTTGTATTGTTGTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAACA 613
Db 701 ATGTTGTATTGTTGTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAACA 760
QY 614 GCTTCTGCCATTTCAAATTTGACCAATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 673
Db 761 GCTTCTGCCATTTCAAATTTGACCAATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 820
QY 674 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 733
Db 821 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 880
QY 734 TTGTGAAGCTGACTATGATGTAATTAATTTCTAAATGAAACGCTGATGAACACAC 793
Db 881 TTGTGAAGCTGACTATGATGTAATTAATTTCTAAATGAAACGCTGATGAACACAC 940
QY 794 TGAATAAACACAGATATCAAAAGTTGTGACTTTTGGAAAGTGCACCGTATGATGAACATC 853
Db 941 TGAATAAACACAGATATCAAAAGTTGTGACTTTTGGAAAGTGCACCGTATGATGAACATC 1000
QY 854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGTT 913
Db 1001 ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGTT 1060
QY 914 GGTCTCAGCTTTGATGTTTGGAGCATAGTGTGATTTCTTATTAATTAATTAATTTGTT 973
Db 1061 GGTCTCAGCTTTGATGTTTGGAGCATAGTGTGATTTCTTATTAATTAATTAATTTGTT 1120
QY 974 TCACAGTCTTTTCACTCATGATGATTAAGAGCACTCGGCAATGATGAACCAATATTAG 1033
Db 1121 TCACAGTCTTTTCACTCATGATGATTAAGAGCACTCGGCAATGATGAACCAATATTAG 1180
QY 1034 GACCCATACCAACACATGATTCAGAAACCAAGAAACGCAAGTATTTTCAACATAACC 1093
Db 1181 GACCCATACCAACACATGATTCAGAAACCAAGAAACGCAAGTATTTTCAACATAACC 1240
QY 1094 AGCTAGATTGGATGAACACACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGCAACCGT 1153
Db 1241 AGCTAGATTGGATGAACACACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGCAACCGT 1300
QY 1154 TGAAGGAATTTATGCTTTTGTATGATGAAGAACATGAGAAACTGTTTGACCTGGTTGAA 1213
Db 1301 TGAAGGAATTTATGCTTTTGTATGATGAAGAACATGAGAAACTGTTTGACCTGGTTGAA 1360
QY 1214 GAATGTTAGATATGATCAACTCAAGAAATTAATCTTGATGAAGCAATGAGCAATCCTT 1273
Db 1361 GAATGTTAGATATGATCAACTCAAGAAATTAATCTTGATGAAGCAATGAGCAATCCTT 1420
QY 1274 TCTTTGACTTTTAAAGAAATGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1333
Db 1421 TCTTTGACTTTTAAAGAAATGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1480
QY 1334 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1393
Db 1481 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1540
QY 1394 AAATTAATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATGATCAATAGCAATTAAC 1453
Db 1541 AAATTAATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATGATCAATAGCAATTAAC 1600
QY 1454 TTGTTAAGCAAGTAGTGTGTTGATATGCTATAGAAATTAATTAATTTTCTTTT 1513
Db 1601 TTGTTAAGCAAGTAGTGTGTTGATATGCTATAGAAATTAATTAATTTTCTTTT 1660

QY 1514 GAAATTACCATTTTAAATATACCTTTGAAATATCCTTTGTGTCAGTGAATATGATG 1573
Db 1661 GAAATTACCATTTTAAATATACCTTTGAAATATCCTTTGTGTCAGTGAATATGATG 1720
QY 1574 ATCTTGCCCTTTTGTACATGGAGGTCACCTCTGAAGTGAATTTTGTGAGTAAAGGAAAT 1633
Db 1721 ATCTTGCCCTTTTGTACATGGAGGTCACCTCTGAAGTGAATTTTGTGAGTAAAGGAAAT 1780
QY 1634 CTTGACTACTTTT 1645
Db 1781 CTTGACTACTTTT 1792

RESULT 8
AAC81772
ID AAC81772 standard; DNA; 1323 BP.
XX
AC AAC81772;
XX
DT 23-FEB-2001 (first entry)
XX
Human beacon partial coding sequence SEQ ID NO: 7.
DE Israeli sand rat; beacon; obesity; NIDDM; energy balance; diabetes;
KW ligand; human; db.
XX
OS Homo sapiens.
XX
PN WO200064931-A1.
XX
PD 02-NOV-2000.
XX
PF 19-APR-2000; 2000WO-AU000342.
XX
PR 23-APR-1999; 99AU-00009919.
XX
PR 24-MAR-2000; 2000AU-00006454.
XX
PA (AUTO-) AUTOGEN PTY LTD.
XX
PI Collier G, Walder K, Zimmet P;
XX
WPI; 2000-687311/67.
XX
Ligand of beacon protein useful for treating obesity, anorexia, energy
imbalance, diabetes, metabolic syndrome, dyslipidemia, hypertension and
insulin resistance.
XX
Claim 4; Fig 2; 67pp; English.
XX
The present invention is related to the isolation of a ligand known as
beacon from the Israeli sand rat. Beacon is associated with the
regulation of energy balance, and the protein, its coding sequence and
analogues can be used in the treatment of diabetes, obesity, anorexia,
energy imbalance, metabolic syndrome, dyslipidaemia, hypertension and
insulin resistance. In addition, they can be used in agriculture to
produce leaner animals
XX
SQ Sequence 1323 BP; 452 A; 199 C; 232 G; 440 T; 0 U; 0 Other;
Query Match 57.7%; Score 1300.6; DB 3; Length 1323;
Best Local Similarity 98.9%; Pred. No. 4.9e-271;
Matches 1309; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 626 TTCAATTTGACCATCAGGAGATGGGTATCAGATCTGCAGTCAATTAATTTTAC 685
Db 1 TTCAATTTGACCATCAGGAGATGGGTATCAGATCTGCAGTCAATTAATTTTAC 60
QY 686 ATCAATAAATTAACCATACAGATCTGAAGCTGAAATATTTTGTGAGAGTCTG 745
Db 61 ATCAATAAATTAACCATACAGATCTGAAGCTGAAATATTTTGTGAGAGTCTG 120
QY 746 ACTATGATGATCAATTAATTTTAAATGAAACGTGATGAACGCACTGAAACACAG 805
Db 746 ACTATGATGATCAATTAATTTTAAATGAAACGTGATGAACGCACTGAAACACAG 805
```

121 ACTATGTAGTCAAATATAATCTTAAATGAAACGATGAACGACACACTGAAAAACACAG 180  
Qy  
806 ATATCAAAAGTTGTTGACATTTTGGAGTGAAGTCAAGTATGATGATGAACATCACAGTACTTTGG 865  
Db  
181 ATATCAAAAGTTGTTGACATTTTGGAGTGAAGTCAAGTATGATGATGAACATCACAGTACTTTGG 240  
Qy  
866 TGTCTACCCGGCAGCTACAGAGCTCCGAGGTCATTTTGGCTTTTAGTGTGGTCTCAGCCCTT 925  
Db  
241 TGTCTACCCGGCAGCTACAGAGCTCCGAGGTCATTTTGGCTTTTAGTGTGGTCTCAGCCCTT 300  
Qy  
926 GTGATGTTTGGAGCATAGTTGTCATTTTATGTAATATTAATCTTGGTTTCAAGTCTTTTC 985  
Db  
301 GTGATGTTTGGAGCATAGTTGTCATTTTATGTAATATTAATCTTGGTTTCAAGTCTTTTC 360  
Qy  
986 AGACTCATGATAGTAAGAGACAGCTGGCAATGATGAGACGATATATAGGACCCATACAC 1045  
Db  
361 AGACTCATGATAGTAAGAGACAGCTGGCAATGATGAGACGATATATAGGACCCATACAC 420  
Qy  
1046 AACACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACCCAGCTAGATTGGG 1105  
Db  
421 AACACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACCCAGCTAGATTGGG 480  
Qy  
1106 ATGAACACAGTCTGCTGGTATGATATGTTAGAGACGCTGCAACCGTTGAAGGAATTTA 1165  
Db  
481 ATGAACACAGTCTGCTGGTATGATATGTTAGAGACGCTGCAACCGTTGAAGGAATTTA 540  
Qy  
1166 TGCTTTGTCATGATGAAGAACATGAGAACTGTTTTCACCTGCTGCGAAGAAATGTTAGAAAT 1225  
Db  
541 TGCTTTGTCATGATGAAGAACATGAGAACTGTTTTCACCTGCTGCGAAGAAATGTTAGAAAT 600  
Qy  
1226 ATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCAGCATCTTTTCTTGACTTTAT 1285  
Db  
601 ATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCAGCATCTTTTCTTGACTTTAT 660  
Qy  
1286 TAAAAAGAAATGAATGGGAATCAAGTGTCTTACTATATATCTTCTAGAGAGATTAAC 1345  
Db  
661 TAAAAAGAAATGAATGGGAATCAAGTGTCTTACTATATATCTTCTAGAGAGATTAAC 720  
Qy  
1346 TTAGACTGTGTCAGTCAACTAAACATCTTAATATTTTGTAAACATTAATTAATTTTGT 1405  
Db  
721 TTAGACTGTGTCAGTCAACTAAACATCTTAATATTTTGTAAACATTAATTAATTTTGT 780  
Qy  
1406 ACAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATTAATTAATTTGTAAGCAAG 1465  
Db  
781 ACAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATTAATTAATTTGTAAGCAAG 840  
Qy  
1466 TATGTCCTTGAATATGCAATAGAAAAATTAATTAATTTTCTTTTGAATTAACCAAT 1525  
Db  
841 TATGTCCTTGAATATGCAATAGAAAAATTAATTAATTTTCTTTTGAATTAACCAAT 900  
Qy  
1526 TTTAAATACCTTTGAAATATCCTTTGTGTCAGTCAATTAATGATGATCTTGGCTTTT 1585  
Db  
901 TTTAAATACCTTTGAAATATCCTTTGTGTCAGTCAATTAATGATGATGATCTTGGCTTTT 960  
Qy  
1586 GTACATGAGGTCACCTCTGAGTGAATTTTTCAGTAAAGAAATCTTGACTTTT 1645  
Db  
961 GTACATGAGGTCACCTCTGAGTGAATTTTTCAGTAAAGAAATCTTGACTTTT 1020  
Qy  
1646 ATATCTTAAAGGAATATCTTTATATATCTTCAATTTAGAACTTAACCTTAAAGAAATTTT 1705  
Db  
1021 ATATCTTAAAGGAATATCTTTATATATCTTCAATTTAGAACTTAACCTTAAAGAAATTTT 1080  
Qy  
1706 TCTTCTGTAATTTGTAACGGGTGATTAATTAATTAATCTAGATAAGCAGGTAAGAAAC 1765  
Db  
1081 TCTTCTGTAATTTGTAACGGGTGATTAATTAATTAATCTAGATAAGCAGGTAAGAAAC 1140  
Qy  
1766 CAAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTGATTTCTTTT 1825  
Db  
1141 CAAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTGTTTCTTTT 1200  
Qy  
1826 TCATTTGGGTGATGTCAGGGTGTATACCCAGACATTCATGGAAGGCAATGTCAGTTTCTCCAT 1885  
Db  
1201 TCATTTGGGTGATGTCAGGGTGTATACCCAGACATTCATGGAAGGCAATGTCAGTTTCTCCAT 1260

1886 TGTGACAGTTTGTTTTAATAAAAAACACATACACACTTTTATTTAAGATTAAAAATCTAAGTGG 1945  
Db  
1261 TGTGACAGTTTGTTTTAATAAAAAACACATACACACTTTTAAAAAATAAAAAA 1320  
Qy  
1946 AAA 1948  
Db  
1321 AAA 1323

RESULT 9  
AD007809  
ID AD007809 standard; cDNA; 1446 BP.  
XX AC AD007809;  
XX 01-JUL-2004 (first entry)  
XX Human polynucleotide #39.  
XX Human; gene; ss; fat cell number; fat cell size; obesity; diabetes;  
KW anorectic; antidiabetic.  
XX Homo sapiens.  
XX US2004071700-A1.  
XX 15-APR-2004.  
XX 09-OCT-2002; 2002US-00267502.  
XX 09-OCT-2002; 2002US-00267502.  
XX (LIFE-) LIFE SCI DEV CORP.  
XX Kim J, Galant R;  
XX WPI; 2004-328526/30.  
XX P-PSDB; ADO08026.  
XX Identifying compounds that influence fat cell number or size for treating  
or preventing obesity or diabetes by exposing the cell to the agent and  
identifying fat cell number or size relative to cells not exposed to the  
agent.  
XX Claim 1; SEQ ID NO 135; 275pp; English.  
XX The invention relates to a method of identifying compounds that influence  
fat cell number or size comprising providing a cell that expresses a gene  
and an agent, exposing the cell to the agent and identifying fat cell  
number or size relative to cells not exposed to the agent. The method  
also comprises providing an expression vector and an agent, exposing the  
vector to the agent, detecting a change in expression of the gene  
relative to expression of the gene in an expression vector not exposed to  
the agent, treating a subject with the agent and identifying fat cell  
number or size in the subject. The agent comprises an antisense  
oligonucleotide. The subject comprises a mammal, preferably a human. The  
method also comprises providing a polypeptide and an agent, exposing the  
polypeptide to the agent, detecting binding of the agent to the  
polypeptide or a change in an activity of the polypeptide, treating a  
subject with the agent and identifying fat cell number or size in the  
subject. The agent comprises an antibody. A method of regulating fat cell  
number or size comprises providing a subject containing fat cells and an  
agent that changes the expression of a gene, and treating the subject  
with the agent under conditions so that fat cell size or number in the  
subject is altered. The method is useful for identifying compounds that  
influence fat cell number or size, for preparing a composition for  
treating or preventing obesity or diabetes. This sequence represents  
human cDNA used in the scope of the invention.  
XX Sequence 1446 BP; 474 A; 258 C; 332 G; 382 T; 0 U; 0 Other;  
SQ

Query Match 57.0%; Score 1286; DB 12; Length 1446;





expressing the protein is useful for producing a pharmaceutical composition. Compositions, antibodies and antisense molecules are useful for the treating a disease associated with STAT6 activation such as allergic diseases, inflammation, autoimmune diseases, diabetes, hyperlipidaemia, infections disease and cancers. Compositions are useful for treating disease associated with STAT6 activation and/or prevention of Th1 hyperactive diseases. Compositions are also useful in rheumatoid arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma, allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage, viral hepatitis and AIDS. The protein has efficient promoting STAT6 activity. The protein or nucleic acid is effectively useful for screening compounds for treating and preventing disease associated with excessive activation or inhibition of STAT6. The present sequence represents a human cDNA encoding a protein which promotes STAT6 activation.

SQ Sequence 1881 BP; 570 A; 344 C; 450 G; 517 T; 0 U; 0 Other;

Query Match      56.6%;    Score 1277.4;    DB 12;    Length 1881;

Best Local Similarity 87.9%; Pred. No. 5.6e-266;

Matches 1413; Conservative 0; Mismatches 192; Indels 3; Gaps 2;

|    |     |  |     |
|----|-----|--|-----|
| Qy | 14  | GTCAATTATTAGAACCAAGTCCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG       | 73  |
| Db | 250 | GTCACTATTATTAGAACCAAGTCCTTGAATGAGAGAGATTATCGGAGCCGGAGATACATTG    | 309 |
| Qy | 74  | ACGAATACAGGAATGACTACTGTGCAAGGATATGTTCTTAGACATTATCACAGAGACATTG    | 133 |
| Db | 310 | ATGAAATACAGAAATGACTACTGCGAAGGATATGTTCCNAGACATTACCATAGAGACGTTG    | 369 |
| Qy | 134 | AAAGCGGGTATCGAATCCAATCGCACTGAGTAAATCTTCAGTCCGACGAGGAAGACGACTCTA  | 193 |
| Db | 370 | AAAGCACTTACCGGATCCATTGCGAGTAAATCCTCAGTCAGGACGAGGAAGACGCCCTA      | 429 |
| Qy | 194 | AAAGGAAGCGAATAGACACTGTGTTCAAGTCATCAGTCACGTTTCGAAGAGCCACCGAAGGA   | 253 |
| Db | 430 | AGAGAAAGCGTAATAGACCCCTGTGCAAGTCATCAGTCGCAATTCGAAGAGCCACCGAAGGA   | 489 |
| Qy | 254 | AAAGATCCAGGAGTATAGAGCATGATGAGGAGGGTCACCTCATCTGTCAAAGTGGAGACG     | 313 |
| Db | 490 | AAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACCTCATCTGTCAAAGTGGAGACG     | 549 |
| Qy | 314 | TTCTTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTTCGCAAAAGTTG | 373 |
| Db | 550 | TTCTTAAGAGCAAGATATGAAATCGTGGACACTTTTAGGTGAAGGAGCCCTTTCGCAAAAGTTG | 609 |
| Qy | 374 | TAGAGTGCATTGATCATGGCATGAGTGCATGTAGCAGTGAAATCGTTAAAAAATG          | 433 |
| Db | 610 | TAGAGTGCATTGATCAGCGCATGAGTGGCTTACATGTAGCAGTGAAAAATGTGAAAAAATG    | 669 |
| Qy | 434 | TAGCCGTTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCACCTTAAATAGTA  | 493 |
| Db | 670 | TAGCCGTTTACCGGAGGCAGCTCGTTCTGAATTCCAAGTATTGGAGCACCTTGAACACA      | 729 |
| Qy | 494 | CTGATCCCAATAGTGTCTTCGAGTGTGTCCAGATGCTAGAAATGTTTGATCATCATGTC      | 553 |
| Db | 730 | CTGACCCCAACAGTGTCTTCGAGTGTCCAGATGCTAGAGTGGTTTTCATCATCATGTC       | 789 |
| Qy | 554 | ATGTTTGTAATGTGTTGTTGAACTACTGCGACCTTAGTACTTACGATTTCAATTAAGAAACA   | 613 |
| Db | 790 | ATGTTTGTAATGTGTTGAGCTGCTGGGACCTTAGTACCTATGATTTATTAAGAAAAATA      | 849 |
| Qy | 614 | GCTTTCTGCCATTTCAAAATTCACACATCAGGCGATGCGGTATCAGATCTGCGCAGTCAA     | 673 |
| Db | 850 | GTTTCTCGCCATTTCAAAATTGATCACATCAGGCAATGCGCTTATCAGATCTGCGCAGTCTA   | 909 |
| Qy | 674 | TAAATTTTTTACATCAATAAATTAACCCCATACAGATCTCGAAGCCTGAAAATATTTTGT     | 733 |
| Db | 910 | TAAATTTTTTACATCAATAAATTAACACACAGGACCTTAAACCTGAAAATATTTTAT        | 969 |
| Qy | 734 | TTGTGAAGCTGACTATGTAGTCAAAATTAATTTCAAATTAAGAACGTGATGAACGCACAC     | 793 |
| Db | 970 | TTGTGAAGCTGACTATGTAGTCAAAATCAATTTCAAATTAAGAACCGAGATGAGCGCACAT    | 102 |



OS Homo sapiens.  
 XX WO200056753-A1.  
 XX  
 XX 28-SEP-2000.  
 XX  
 XX 16-MAR-2000; 2000WO-US006765.  
 XX  
 XX 23-MAR-1999; 99US-0126051P.  
 PR 10-DEC-1999; 99US-0169906P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Ruben SM, Komatsoulis G;  
 XX WPI; 2000-594570/56.  
 DR P-PSDB; AAB33724.  
 XX  
 XX Nucleic acid molecules encoding human secreted proteins, used in  
 PT preventing, treating or ameliorating a disorder.  
 XX  
 PS Claim 1; Page 349-350; 410pp; English.  
 XX  
 CC The invention relates to the isolation of genes AAC59277-C59325 encoding  
 CC 49 human secreted proteins AAB33718-B33764. The genes can be used to  
 CC generate fusion proteins by linking to the gene for the human  
 CC immunoglobulin G Fc portion (SEQID1) for increasing the stability of the  
 CC fusion protein as compared to the human protein only. The genes and  
 CC proteins are useful for preventing, ameliorating or treating medical  
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from  
 CC a range of human tissues disclosed in the specification. The nucleic  
 CC acids, proteins, antibodies and (ant)agonists are useful in the  
 CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)  
 CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections  
 XX  
 SQ Sequence 1222 BP; 407 A; 194 C; 229 G; 392 T; 0 U; 0 Other;  
 Query Match 52.5%; Score 1185; DB 3; Length 1222;  
 Best Local Similarity 99.9%; Pred No. 4.6e-246;  
 Matches 1196; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 446 GTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTACTGATCCCAATA 505  
 9 GTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTACTGATCCCAATA 68  
 506 GTGCTTCGGATGTCAGATGCTAGATGGTTGATCATCATGCTGATGTTGTATTG 565  
 69 GTGCTTCGGATGTCAGATGCTAGATGGTTGATCATCATGCTGATGTTGTATTG 128  
 566 TGTTCGAACTCTGGGACTTTAGTACTTACGATTTTCATTAAGAAACAGCTTCTGCGCAT 625  
 129 TGTTCGAACTCTGGGACTTTAGTACTTACGATTTTCATTAAGAAACAGCTTCTGCGCAT 188  
 626 TTCAAATTTGACCATCATCAGGATGCGGTATCAGATCTGCCAGTCAATAAATTTTAC 685  
 189 TTCAAATTTGACCATCATCAGGATGCGGTATCAGATCTGCCAGTCAATAAATTTTAC 248  
 686 ATCATAAATAAATTAACCATCAGATCTGAGCCCTGAAATATTTTGTGTAAGTCTG 745  
 249 ATCATAAATAAATTAACCATCAGATCTGAGCCCTGAAATATTTTGTGTAAGTCTG 308  
 746 ACTATGATGCTAAATATTTCTAAATGAACGATGATGACGACACTGAAACACAG 805  
 309 ACTATGATGCTAAATATTTCTAAATGAACGATGATGACGACACTGAAACACAG 368

QY 806 ATATCAAAAGTTGTTGACCTTTGGAAGTGCACGATGATGATGAACATCACAGTACTTTGG 865  
 DB 369 ATATCAAAAGTTGTTGACCTTTGGAAGTGCACGATGATGATGAACATCACAGTACTTTGG 428  
 QY 866 TGTCTACCCGGCACTACAGAGCTCCCGAGGTCATTTTGGCTTTAGTTGGTCTCAGCCTT 925  
 DB 429 TGTCTACCCGGCACTACAGAGCTCCCGAGGTCATTTTGGCTTTAGTTGGTCTCAGCCTT 488  
 QY 926 GTGATGTTTGGAGCATAGTTGCAATCTTATTTGAATATTTACCTTGGTTTCAAGTCTTTC 985  
 DB 489 GTGATGTTTGGAGCATAGTTGCAATCTTATTTGAATATTTACCTTGGTTTCAAGTCTTTC 548  
 QY 986 AGACTCATGATAGTAAGAGCACCTTGGCAATGATGGAACGAAATATTAGGACCCATACCAC 1045  
 DB 549 AGACTCATGATAGTAAGAGCACCTTGGCAATGATGGAACGAAATATTAGGACCCATACCAC 608  
 QY 1046 AACACATGATTCAGAAAACAAAGAAAACGCAAGTATTTTTCACATACACAGCTAGATTGGG 1105  
 DB 609 AACACATGATTCAGAAAACAAAGAAAACGCAAGTATTTTTCACATACACAGCTAGATTGGG 668  
 QY 1106 ATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTA 1165  
 DB 669 ATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTA 728  
 QY 1166 TGTCTTGTGATGATGAAGAACATGAGAAACCTGTTTACCTGGTTGCAAGAAATGTTAGAT 1225  
 DB 729 TGTCTTGTGATGATGAAGAACATGAGAAACCTGTTTACCTGGTTGCAAGAAATGTTAGAT 788  
 QY 1226 ATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTCAGCATCTCTTCTTTGACTTAT 1285  
 DB 789 ATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTCAGCATCTCTTCTTTGACTTAT 848  
 QY 1286 TAAAAAGAAATGAATGGGAATCAGTGGTCTTACTATATATATCTTCTTAGAAGAGATTAC 1345  
 DB 849 TAAAAAGAAATGAATGGGAATCAGTGGTCTTACTATATATCTTCTTAGAAGAGATTAC 908  
 QY 1346 TTAAGACTGTGTCAGTCACTAAACATCTTAATATTTTGTAAACATTAATATTTTGT 1405  
 DB 909 TTAAGACTGTGTCAGTCACTAAACATCTTAATATTTTGTAAACATTAATATTTTGT 968  
 QY 1406 ACAGTTAAGTCTAAATATTTGATGTTTGTATCAATAGCAATTAATTAATCTGTTAAGCAAG 1465  
 DB 969 ACAGTTAAGTCTAAATATTTGATGTTTGTATCAATAGCAATTAATTAATCTGTTAAGCAAG 1028  
 QY 1466 TATGGTCTTGTATAATGCAATTAAGAAAAATTAATTTTCTTTTGTAAATTAACCAT 1525  
 DB 1029 TATGGTCTTGTATAATGCAATTAAGAAAAATTAATTTTCTTTTGTAAATTAACCAT 1088  
 QY 1526 TTTAAATACCTTTGAAATAT-CCCTTGTGTCAGTGAATAATGATGATCTTGCCTTT 1584  
 DB 1089 TTTAAATACCTTTGAAATATATACCTTTGTGTCAGTGAATAATGATGATCTTGCCTTT 1148  
 QY 1585 TGTACATGGAGGTCACCTCTCAAGTATTTTGTGATTAAGAAAGAAATCTTGACTA 1641  
 DB 1149 TGTACATGGAGGTCACCTCTCAAGTATTTTGTGATTAAGAAAGAAATCTTGACTA 1205  
 RESULT 12  
 ADI31373  
 ID ADI31373 standard; cDNA; 1456 BP.  
 XX  
 AC ADI31373;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Human cDNA #699.  
 XX  
 KW Human; gene; ss; immunological response; immunopathological condition;  
 KW Crohn's disease; asthma; ulcerative colitis; hyper eosinophilia;  
 KW irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;  
 KW acute monocytic leukaemia; antiinflammatory; antisthmatic; antiulcer;  
 KW osteopathic; antiarthritic; antifurheumatic; cytostatic.  
 XX



DE Human cDNA encoding protein that promotes STAT6 activation #55.  
 XX  
 KW ss; gene; human; signal transducer and activator of transcription 6;  
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;  
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;  
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;  
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;  
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003104277-A2.  
 XX  
 PD 18-DEC-2003.  
 XX  
 PF 05-JUN-2003; 2003WO-JP007123.  
 XX  
 XX 05-JUN-2002; 2002JP-00164257.  
 PR 06-JUN-2002; 2002US-0385912P.  
 PR 26-DEC-2002; 2002JP-00377326.  
 PR 27-DEC-2002; 2002US-0436467P.  
 PR 15-MAY-2003; 2003JP-00137505.  
 PR 16-MAY-2003; 2003US-0470836P.  
 XX  
 XX (ASAH ) ASAMI KASEI KK.  
 PA  
 PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;  
 XX  
 XX WPI: 2004-122214/12.  
 DR P-PSDB; ADI26145.  
 DR  
 XX  
 PT New signal transducer and activator of transcription 6 activation  
 PT promoting purified protein, for diagnosing and treating disease  
 PT associated with activation/inhibition of transcription factor e.g.  
 PT diabetes and cancer.  
 PT  
 XX  
 PS Claim 4; SEQ ID NO 109; 1368pp; English.  
 XX  
 CC The invention relates to a purified protein promoting signal transducer  
 CC and activator of transcription 6 activation (STAT6). The protein is  
 CC useful for the producing an antibody, which involves administering the  
 CC protein or its epitope-bearing fragments to a non-human animal as an  
 CC antigen. The nucleic acid is useful for diagnosing a disease or  
 CC susceptibility to a disease related to expression or activity of the  
 CC protein. A transformant expressing the protein is useful for screening  
 CC compounds which inhibit or promote STAT6 activation. A transformant  
 CC expressing the protein is useful for producing a pharmaceutical  
 CC composition. Compositions, antibodies and antisense molecules are useful  
 CC for the treating a disease associated with STAT6 activation such as  
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,  
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful  
 CC for treating disease associated with STAT6 activation and/or prevention  
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid  
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,  
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,  
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6  
 CC activity. The protein or nucleic acid is effectively useful for screening  
 CC compounds for treating and preventing disease associated with excessive  
 CC activation or inhibition of STAT6. The present sequence represents a  
 CC human cDNA encoding a protein which promotes STAT6 activation.  
 XX  
 SQ Sequence 1865 BP; 629 A; 309 C; 381 G; 546 T; 0 U; 0 Other;  
 Query Match 49.6%; Score 1118.8; DB 12; Length 1865;  
 Best Local Similarity 90.2%; Pred. No. 1e-231;  
 Matches 1191; Conservative 0; Mismatches 129; Indels 0; Gaps 0;  
 Query 329 ATGAATCGTGACACTTTGGGTGAAGGAGCTTTGGCAAGTTGTAGAGTGCATTGATC 388  
 DB 545 ATGAATTTGTTGATCTTTAGTGAAGGAGCTTTGGCAAGTTGTAGAGTGCATTGATC 604  
 Query 389 ATGGCATGGATGGCATGTAGCAGTGAAGTAAATCGTAAATAATGAGCCGTTACCGTG 448

605 ATAAAGCGGAGGTAGACATGTAGCAGTAAATAATAGTTAAATAATGTGGATAGATCTGTG 664  
 QY  
 449 AAGCAGCTCGTTCAAGAAATCCAAGTATTAGAGCAGCTTAAATAGTACTGTATCCCAATAGTG 508  
 DB  
 665 AAGCTGCTCGTCAAGAAATCAAGTTCTTGAACATCTGAATACAAACAGACCCCAACAGTA 724  
 QY  
 509 TCTCCGATGTCTCAGATGCTAGAGTGGTTTGAATCATCATGCTCATGTTTGTATGTGT 568  
 DB  
 725 CTTTCGCTGTCTCAGATGTTGGAATGTTTGAAGTATGAGCATCATGTCATCTTGCATTTGTT 784  
 QY  
 569 TTGAATCTAGTGGACTTTAGTACTTACGATTTTCAATTAAGAAACAGCTTTTCCCATTTTC 628  
 DB  
 785 TTGAATCTAGTGGACTTTAGTACTTACGATTTTCAATTAAGAAATGCTTTTACCATTTTC 844  
 QY  
 629 AAATTGACCATCATCAGGAGATGGCTATCAGATCTGCCAGTCAATAAATTTTTTACATC 688  
 DB  
 845 GACTGATCATATCAGAAAGATGSCATATCAGATGCAAGTCTCTGAAATTTTTTGCACA 904  
 QY  
 689 ATATAAATTAACCATACAGATCTGAAGCCCTGAAAAATATTTTGTGTGAAGTCTGACT 748  
 DB  
 905 GTAATAAGTTGACTCACACAGACTTTAAAGCCCTGAAAAACATCTTATTTGTGCACTGCT 964  
 QY  
 749 ATGTAGTCAAAATATAATTTCTAAATGAAACGTTGATGAACGACACCTGAAAAACAGATA 808  
 DB  
 965 ACACAGAGCGGTATTAATCCCAAAATAAACGTTGATGAACGACCTTAAATAATCCAGATA 1024  
 QY  
 809 TCAAAGTTTGTGACTTTTGGAAAGTCAACGCTATGATGATGATCAACATCACAGTACTTTGGTGT 868  
 DB  
 1025 TTAAGTTGTAGACTTTTGTAGTCAACATATGATGACGACATCACAGTACATTTGTTAT 1084  
 QY  
 869 CTACCCGCACTACAGAGCTCCGAGGTCATTTTGGCTTTTAGTGGTCTCAGCCTTGTG 928  
 DB  
 1085 CTCAAGACATTTATAGAGCAGCTGAAAGTTATTTTAGCCCTAGGGTGGTCCACAGCCTTGTG 1144  
 QY  
 929 ATGTTTGAGCATAGTTGTCATCTTATTTGATATTTACCTTGGTTTTCAGCTCTTCAGA 988  
 DB  
 1145 ATGTTTGAGCATAGTTGTCATCTTATTAATGAATATTACCTTGGTTTTCAGCTCTTCAGA 1204  
 QY  
 989 CTCATGATGATGAAGAGCAGCTGGCAATGATGATGAAGCAATATTAGGACCCATACCAAC 1048  
 DB  
 1205 CTCATGATGATGAAGAGCAGCTGGCAATGATGATGAAGCAATATTAGGACCCATACCAAC 1264  
 QY  
 1049 ACATGATTCAGAAAACAAAGAAAACGCAAGTATTTTCAACATAACAGCTAGATTGGGATG 1108  
 DB  
 1265 ACATGATTCAGAAAACAAAGAAAACGCAAGTATTTTCAACATAACAGCTAGATTGGGATG 1324  
 QY  
 1109 AACACAGTTCTGCTGGTATGATGTTAGGAGCGCTGCAACCCGTTGAAGGAATTTATGC 1168  
 DB  
 1325 AACACAGTTCTGCTGGTATGATGTTAGGAGCGCTGCAACCCGTTGAAGGAATTTATGC 1384  
 QY  
 1169 TTTGTCATGATGAAGAACATCAGAAACCTGTTTGAACCTGTTTGAAGGAATTTATGC 1228  
 DB  
 1385 TTTGTCATGATGAAGAACATCAGAAACCTGTTTGAACCTGTTTGAAGGAATTTATGC 1444  
 QY  
 1229 ATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCAGCAGCTCTTTTCTTGTACTTAA 1288  
 DB  
 1445 ATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCAGCAGCTCTTTTCTTGTACTTAA 1504  
 QY  
 1289 AAAGAAATGAATGGGAATCAGTGGTCTTACTATATATCTCTAGAGAGATTAATTA 1348  
 DB  
 1505 AAAGAAATGAATGGGAATCAGTGGTCTTACTATATATCTCTAGAGAGATTAATTA 1564  
 QY  
 1349 AGACTGTGTCAGTCAACTAAACATTTCTAATAATTTTGAACATTAATAATTTTGTACA 1408  
 DB  
 1565 AGACTGTGTCAGTCAACTAAACATTTCTAATAATTTTGAACATTAATAATTTTGTACA 1624  
 QY  
 1409 GTTAAGTGTAAATATGTATGTTTGTATCAATAGCAATAATTAATCTGTTAAGCAAGTAT 1468  
 DB  
 1625 GTTAAGTGTAAATATGTATGTTTGTATCAATAGCAATAATTAATCTGTTAAGCAAGTAT 1684  
 QY  
 1469 GGTCTTGTATATGCAATTAAGAAAAATTAATAATTTTCTTTTGTAAATTAACCAATTTT 1528  
 DB  
 1685 GGTCTTGTATATGCAATTAAGAAAAATTAATAATTTTCTTTTGTAAATTAACCAATTTT 1744

QY 1529 AAATACCTTTGAAATATATCTTTGTCAGTGATAAAATGATGATTCATCTTCCCTTTTGTA 1598  
DB 1745 AAATACCTTTGAAATATATCTTTGTCAGTGATAAAATGATGATTCATCTTCCCTTTTGTA 1804  
QY 1589 CATGAGGTCACTCTGAGTGAATTTTTTTGAGTAAAGGAAATCTTGATCTATTATA 1648  
DB 1805 CATGAGGTCACTCTGAGTGAATTTTTTTGAGTAAAGGAAATCTTGATCTATTATA 1864

RESULT 14  
ID ADO07812  
XX ADO07812 standard; cDNA; 1446 BP.  
AC ADO07812;  
XX 01-JUL-2004 (first entry)  
DT Mouse polynucleotide #36.  
DE Mouse; gene; ss; fat cell number; fat cell size; obesity; diabetes;  
KW anorectic; antidiabetic.  
KW Mus sp.  
XX US2004071700-A1.  
XX 15-APR-2004.  
PD 09-OCT-2002; 2002US-00267502.  
PF 09-OCT-2002; 2002US-00267502.  
PR (LIFE-) LIFE SCI DEV CORP.  
XX Kim J, Galant R;  
XX WPI; 2004-328526/30.  
DR P-PSDB; ADO08029.  
XX Identifying compounds that influence fat cell number or size for treating  
PT or preventing obesity or diabetes by exposing the cell to the agent and  
PT identifying fat cell number or size relative to cells not exposed to the  
PT agent.  
XX Claim 1; SEQ ID NO 138; 275pp; English.

XX The invention relates to a method of identifying compounds that influence  
CC fat cell number or size comprising providing a cell that expresses a gene  
CC and an agent, exposing the cell to the agent and identifying fat cell  
CC number or size relative to cells not exposed to the agent. The method  
CC also comprises providing an expression vector and an agent, exposing the  
CC vector to the agent, detecting a change in expression of the gene  
CC relative to expression of the gene in an expression vector not exposed to  
CC the agent, treating a subject with the agent and identifying fat cell  
CC number or size in the subject. The agent comprises an antisense  
CC oligonucleotide. The subject comprises a mammal, preferably a human. The  
CC method also comprises providing a polypeptide and an agent, exposing the  
CC polypeptide to the agent, detecting binding of the agent to the  
CC polypeptide or a change in an activity of the polypeptide, treating a  
CC subject with the agent and identifying fat cell number or size in the  
CC subject. The agent comprises an antibody. A method of regulating fat cell  
CC number or size comprises providing a subject containing fat cells and an  
CC agent that changes the expression of a gene, and treating the subject  
CC with the agent under conditions so that fat cell size or number in the  
CC subject is altered. The method is useful for identifying compounds that  
CC influence fat cell number or size, for preparing a composition for  
CC treating or preventing obesity or diabetes. This sequence represents  
CC mouse cDNA used in the scope of the invention.

XX Sequence 1446 BP; 461 A; 272 C; 352 G; 361 T; 0 U; 0 Other;

Query Match 48.6%; Score 1097.2; DB 12; Length 1446;

Best Local Similarity 90.8%; Pred. No. 4.5e-227;  
Matches 1168; Conservative 0; Mismatches 118; Indels 0; Gaps 0;  
QY 14 GTCAATTTAGAACCAAGGTCCTTTGAATGAGCGAGATTAATCGGGACCGGAGATACGTTG 73  
DB 161 GTCATATTTAGAAAGCAAGATGCTTTGAATGAGAGATTAATCGGGACCGGAGATACATG 220  
QY 74 ACGAATACAGGAATGACTTACTGTGAAGGATATGTTCTCTAGACATTAATACAGAGACATG 133  
DB 221 ATGAATACAGAAATGACTTACTGTGCAAGGATATGTTTCCAAGACATTAATCATAGACGTTG 280  
QY 134 AAAGCGGTATCGAATTCATCGAGTAAATCTTCACTCCGCGAGGAGGAGGAGTCTTA 193  
DB 281 AAAGCACTTACCGGATCCATTCGAGTAAATCTTCACTCCGCGAGGAGGAGGAGTCTTA 340  
QY 194 AAAGCAAGCGCAATAGACACATGTTCAAGTCAATCAGTCACTGTTCAAGAGCCACCAAGGA 253  
DB 341 AGAAGAGCGTAATAGACCTTGTGCAAGTCAATCAGTGCATTCAGAGGCCACCGAAGGA 400  
QY 254 AAAGTCCAGGATATAGAGGATGATGAGGAGGTCACTCTGATCTGTCAAAAGTGAGAGC 313  
DB 401 AAAGTCCAGGATATAGAGGATGATGAGGAGGTCACTCTGATCTGTCAAAAGTGAGAGC 460  
QY 314 TTCTAAGAGCAAGATATGAATCTGTGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 373  
DB 461 TTCTAAGAGCAAGATATGAATCTGTGACACTTTTAGGTGAAGGAGCCCTTTGGCAAGTTG 520  
QY 374 TAGAGTGCATTTGATCATGGCATGATGCGATGATGAGTGAAGGAGTGAAGGAGTGAAGGAGT 433  
DB 521 TAGAGTGCATTTGATCATGGCATGATGCGATGATGAGTGAAGGAGTGAAGGAGTGAAGGAGT 580  
QY 434 TAGGCGGTTTACCGTGAAGGAGTCTGTTTCAAGAAATCCAAGTATTAAGAGCACCTTAATAGTA 493  
DB 581 TAGGAGCGTTACCGGAGGAGTCTGTTTCAAGAAATCCAAGTATTAAGAGCACCTTAATAGTA 640  
QY 494 CTGATCCCAATAGTGTCTTCGATGTCGATGATGCTAGATGCTAGATGCTTGTGATCATGCTC 553  
DB 641 CTGACCCCAACAGTGTCTTCGATGTCGATGATGCTAGATGCTTGTGATCATGCTC 700  
QY 554 ATGTTTGTATTTGTTGAACTTCTGGGACTTAGTACTTACGATTTTCAATTAAGAAAAACA 613  
DB 701 ATGTTTGTATTTGTTGAGGCTGCTGGGACTTAGTACTTATTAATTAAGAAAAACA 760  
QY 614 GCTTTCTGCCATTTCAAATTCACCATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 673  
DB 761 GTTTCTGCCATTTCAAATTCACCATCAGGAGTAAATGGCTTATCAGATCTGCCAGTCTA 820  
QY 674 TAAATTTTACATCATTAATAAATTAACCCATACAGATCTCGAAGCTGAAATATTTTGT 733  
DB 821 TAAATTTTACATCATTAATAAATTAACCCATACAGATCTCGAAGCTGAAATATTTTAT 880  
QY 734 TTGTGAAGTCTGATCTAGTCAAAATATAATTTCTAAATGAAACGCTGATGAACGCACAC 793  
DB 881 TTGTGAAGTCTGATCTAGTCAAAATATAATTTCTAAATGAAACGCTGATGAACGCACAC 940  
QY 794 TGAAGAAACACAGATATCAAAGTGTGTTGCTTTGGAAGTGAACGCTGATGAATGAACATC 853  
DB 941 TGAAGAAACACAGATATCAAAGTGTGTTGCTTTTGAAGTGAACATATCAGCAGCAACATC 1000  
QY 854 ACAGTACTTTGGTGTCTACCCGGGACATCAGAGCTCCCGAGGTCATTTGGCTTTAGTT 913  
DB 1001 ATAGTACTTTGGTGTCTCAAGGGGACATCAGGGGCTCCAGAGGTCATTTTGGCTCTAGTT 1060  
QY 914 GGTCTCAGCCTTTGTGATGTTTGGAGCATAGTGTGATTTCTTATTTGAAATATTACCTTGGTT 973  
DB 1061 GGTCTCAGCCTTTGTGATGTTTGGAGCATAGTGTGATTTCTTATTTAGTACTACCTTGGT 1120  
QY 974 TCACAGTCTTTTCAAGTCAATGATAGTAAAGAGCCTCGGCAATGATGAACATATTAG 1033  
DB 1121 TCACAGTCTTTTCAAGTCAATGATAGTAAAGAGCCTCGGCAATGATGAAGGAGTCTTAG 1180  
QY 1034 GACCCATACCAACACATGATTCAGAAACAGAAACGCAAGTATTTTCCACCAATAACC 1093

```
Db      1181 GACCCATCCAGCACATATGATCCAGAGACAGGAAACGCAAGTATTTCACCATACC 1240
Qy      1094 ACCTAGATTGGATGACACAGCTTCCTGCTGGTAGATATGTTAGGAGACGCTGCAACCGT 1153
Db      1241 AGCTAGATTGGAGCAGCAGCATAGTTCAGCTGGGAGATATGTTAGGAGACGCTGCAACCGT 1300
Qy      1154 TGAAGGATTTATGCTTTGTCATGATGAGAACATGAGAACTGTTGACCTGGTTCGAA 1213
Db      1301 TAAAGGAATTTATGCTGTGTCATGACGAAGACATGAGAAAGCTGTTGACCTGGTTCGAA 1360
Qy      1214 GAATGTTAGATATGATCCAACTCAAAGAAATTACCTTGGATCAAGCATTCGAGCATCCTT 1273
Db      1361 GAATGTTGAGTATGACCCAGCAGGAGAGATCACCCTGGATGAGCATTCGAGCACCTT 1420
Qy      1274 TCTTTGACTTATTAATAAAGAAATGA 1299
Db      1421 TCTTTGACTTATTAATAAAGAAATGA 1446

RESULT 15
AAD32039
ID      AAD32039 standard; DNA; 21234 BP.
XX
AC      AAD32039;
XX
DT      18-JUN-2002 (first entry)
XX
DE      Human kinase protein gene.
XX
KW      Human; kinase protein; enzyme; cytostatic; osteopathic; gene expression;
KW      colon-moderately differentiated adenocarcinoma; chromosome mapping; gene;
KW      drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;
KW      bone osteosarcoma; single nucleotide polymorphism; SNP; ds.
XX
OS      Homo sapiens.
XX
FH      Key
FT      variation
FT      Location/Qualifiers
FT      replace(75..76,A)
FT      /*tag= a
FT      /standard_name= "Single nucleotide polymorphism"
FT      exon
FT      2007..2059
FT      /*tag= b
FT      intron
FT      2060..3118
FT      /*tag= c
FT      exon
FT      3119..3341
FT      /*tag= d
FT      intron
FT      3342..4462
FT      /*tag= e
FT      exon
FT      4463..4553
FT      /*tag= f
FT      intron
FT      4554..4948
FT      /*tag= g
FT      exon
FT      4949..5015
FT      /*tag= h
FT      intron
FT      5016..8054
FT      /*tag= i
FT      variation
FT      replace(7980,C)
FT      /*tag= j
FT      /standard_name= "Single nucleotide polymorphism"
FT      exon
FT      8055..8171
FT      /*tag= k
FT      intron
FT      8172..8258
FT      /*tag= l
FT      exon
FT      8259..8425
FT      /*tag= m
FT      intron
FT      8426..9007
FT      /*tag= n
FT      variation
FT      replace(8571,T)
FT      /*tag= o
FT      /standard_name= "Single nucleotide polymorphism"
FT      exon
FT      9008..9102
FT      /*tag= p
FT      intron
FT      9103..9352
```

```
FT      exon
FT      /*tag= q
FT      9353..9482
FT      /*tag= r
FT      intron
FT      9483..13437
FT      /*tag= s
FT      variation
FT      replace(11257,Y)
FT      /*tag= t
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(11684,C)
FT      /*tag= u
FT      variation
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(13312,C)
FT      /*tag= v
FT      /standard_name= "Single nucleotide polymorphism"
FT      13438..13520
FT      /*tag= w
FT      intron
FT      13521..16514
FT      /*tag= x
FT      exon
FT      16515..16594
FT      /*tag= y
FT      intron
FT      16595..17962
FT      /*tag= z
FT      variation
FT      replace(17110,T)
FT      /*tag= aa
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(17451,C)
FT      /*tag= ab
FT      variation
FT      17963..18053
FT      /*tag= ac
FT      intron
FT      18054..18139
FT      /*tag= ad
FT      exon
FT      18140..18280
FT      /*tag= ae
FT      variation
FT      replace(20766,G)
FT      /*tag= af
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(20913..20914,T)
FT      /*tag= ag
FT      /standard_name= "Single nucleotide polymorphism"
FT      WO200216567-A2.
XX
PN      28-FEB-2002.
XX
PD      24-AUG-2001; 2001WO-US026389.
XX
PF      24-AUG-2000; 2000US-0227470P.
XX
PR      19-MAR-2001; 2001US-00810671.
XX
PA      (APPL-) APPLERA CORP.
XX
PI      Yan C, Ye J, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR      WPI; 2002-269354/31.
XX
PT      New human kinase proteins and nucleic acids, useful in drug screening
PT      assays, identifying modulators of kinase activity or treating disorders
PT      characterized by absence or unwanted expression of the protein.
XX
PS      Claim 4; Fig 3; 8lpp; English.
XX
CC      The invention relates to isolated human kinase proteins and nucleic
CC      acids. The nucleic acid and peptide sequences can be used as models for
CC      the development of human therapeutic targets, aid in the identification
CC      of therapeutic proteins and serve as targets for the development of human
CC      therapeutic agents that modulate kinase activity in cells and tissues
CC      that express the kinase. The nucleic acids are useful as probes or
CC      primers, in constructing recombinant vectors, for expressing antigenic
CC      portions of the proteins, chromosome mapping, drug screening, testing an
CC      individual for a genotype, and for gene therapy in patients containing
CC      cells that are aberrant in kinase gene expression. The proteins may be
CC      used in drug screening assays, in the identification of compounds that
```

modulate, stimulate or inhibit kinase activity, in pharmacogenomic analysis, in treating disorders characterised by an absence or unwanted expression of the protein (bone osteosarcoma, or colon-moderately differentiated adenocarcinoma), and in generating antibodies specific for the peptides. Such antibodies can be used to detect the protein in situ, in vitro, or in cell lysate or supernatant, to isolate and purify the proteins from host cells, pharmacogenomic analysis, tissue typing, and in inhibiting protein function. The present sequence is human kinase protein gene located on chromosome 5

CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
XX  
SQ

Sequence 21234 BP; 6065 A; 3772 C; 4258 G; 7139 T; 0 U; 0 Other;

Query Match 48.6%; Score 1097; DB 6; Length 21234;  
Best Local Similarity 100.0%; Pred. No. 1e-226;  
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1157 AGGAATTTATGCTTGTGATGATGAAGAAACATGAGAACTGTTGACCTGGTTCGAAGAA 1216  
Db 18138 AGGAATTTATGCTTGTGATGATGAAGAAACATGAGAACTGTTGACCTGGTTCGAAGAA 18197

Qy 1217 TGTTAGATATGATCCAACTCAAGAAATTTACCTGGATGAAGCAATGCGAGCATCCTTCT 1276  
Db 18198 TGTTAGATATGATCCAACTCAAGAAATTTACCTGGATGAAGCAATGCGAGCATCCTTCT 18257

Qy 1277 TTGACTTATTAATAAAGAAATGAAATGGGAATCAAGTGGTCTTACTATATATCTTCTCTAGA 1336  
Db 18258 TTGACTTATTAATAAAGAAATGAAATGGGAATCAAGTGGTCTTACTATATATCTTCTCTAGA 18317

Qy 1337 AGAGATTAATTAAGCTGTGTCAGTCAACTAAACATTTCTATATATTTTGTAAACATTAATA 1396  
Db 18318 AGAGATTAATTAAGCTGTGTCAGTCAACTAAACATTTCTATATATTTTGTAAACATTAATA 18377

Qy 1397 TTATTTGTGACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATATTAATTAACCTTG 1456  
Db 18378 TTATTTGTGACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATATTAATTAACCTTG 18437

Qy 1457 TTAAGCAAGTATGGCTTTGATAATGATTAAGTGAATTAATAATTAATTTCTTTTGTAA 1516  
Db 18438 TTAAGCAAGTATGGCTTTGATAATGATTAAGTGAATTAATAATTAATTTCTTTTGTAA 18497

Qy 1517 ATTACCAATTTTAATACCTTTGAAATATCTTTGTGTCAGTGAATAATGATGATGATC 1576  
Db 18498 ATTACCAATTTTAATACCTTTGAAATATCTTTGTGTCAGTGAATAATGATGATGATC 18557

Qy 1577 TTGCTTTTGTACATGGAGTCACTCTGAAGTGAATTTTGTAGTAAAGGAATCTT 1636  
Db 18558 TTGCTTTTGTACATGGAGTCACTCTGAAGTGAATTTTGTAGTAAAGGAATCTT 18617

Qy 1637 GACTACTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTAGAACTTAACCTTT 1696  
Db 18618 GACTACTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTAGAACTTAACCTTT 18677

Qy 1697 AAAAGTTTCTTCTGTAATTTGTGAACGGGTGATTTATTAATTAATCTAGATAAGCAGGT 1756  
Db 18678 AAAAGTTTCTTCTGTAATTTGTGAACGGGTGATTTATTAATTAATCTAGATAAGCAGGT 18737

Qy 1757 ACTAGAAACCAAACTCAGAAATCTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTG 1816  
Db 18738 ACTAGAAACCAAACTCAGAAATCTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTG 18797

Qy 1817 TATTTCTTTTCAATGGGTGATGTCAGGGTGAATTAACCAATTCATGAAAGGCATGCGAG 1876  
Db 18798 TATTTCTTTTCAATGGGTGATGTCAGGGTGAATTAACCAATTCATGAAAGGCATGCGAG 18857

Qy 1877 TTTGTCCATTTGTGACAGTTTGTATTAATAAACCAATACACACTTTTATTAAGATTAATA 1936  
Db 18858 TTTGTCCATTTGTGACAGTTTGTATTAATAAACCAATACACACTTTTATTAAGATTAATA 18917

Qy 1937 TCTAACTGGAAGTCAAGCTTGGAAATGCAATTTCCAAAGTATGTTTGGTGAAGTCAACAGA 1996  
Db 18918 TCTAACTGGAAGTCAAGCTTGGAAATGCAATTTCCAAAGTATGTTTGGTGAAGTCAACAGA 18977

Qy 1997 TATAAAAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTAATACCAAGTCTTTAGGAGTCT 2056

Search completed: March 16, 2005, 09:01:21  
Job time : 1200.6 secs

Db 18978 TATAAAAATAGAAATTTCTGATGAGAGTTTCAGTTTTTAATACCAAGTCTTTAGAGTCT 19037  
Qy 2057 TAACATTTGGCCAGCATCTGTTTATCAAAATGACATATAATACGTAACCTTATAAGAAATTAAG 2116  
Db 19038 TAACATTTGGCCAGCATCTGTTTATCAAAATGACATATAATACGTAACCTTATAAGAAATTAAG 19097  
Qy 2117 TTTTATTAATTAGGCAATTTATGCTGTGATAATTTCTTACGGGAGAAAGAGGATTTGATTG 2176  
Db 19098 TTTTATTAATTAGGCAATTTATGCTGTGATAATTTCTTACGGGAGAAAGAGGATTTGATTG 19157

Qy 2177 GAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAATGATTTGCTTACAT 2236  
Db 19158 GAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAATGATTTGCTTACAT 19217

Qy 2237 AAACCTTTTGTACTTTCAG 2253  
Db 19218 AAACCTTTTGTACTTTCAG 19234

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 04:53:38 ; Search time 9703.74 Seconds  
(without alignments)  
11265.245 Million cell updates/sec

Title: US-10-801-671-1\_COPY\_72\_2327

Perfect score: 2256  
Sequence: 1 gacactcatccagtcattt.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_ats.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID       | Description        |
|------------|--------|-------------|--------|-------------|--------------------|
| 1          | 2256   | 100.0       | 2354   | 6 AR232168  | AR232168 Sequence  |
| 2          | 2256   | 100.0       | 2354   | 6 AR406007  | AR406007 Sequence  |
| 3          | 2256   | 100.0       | 2354   | 6 AR534243  | AR534243 Sequence  |
| 4          | 2256   | 100.0       | 2354   | 6 AR392827  | AR392827 Sequence  |
| 5          | 2048   | 90.8        | 2488   | 6 AX056410  | AX056410 Sequence  |
| 6          | 2046.6 | 90.7        | 2446   | 6 AR077239  | AR077239 Sequence  |
| 7          | 2046.6 | 90.7        | 2446   | 6 BD107903  | BD107903 Protein k |
| 8          | 1632   | 72.3        | 1814   | 9 AF294429  | AF294429 Homo sapi |
| 9          | 1629.4 | 72.2        | 1792   | 6 AX961896  | AX961896 Sequence  |
| 10         | 1318.4 | 58.4        | 3492   | 9 BC063116  | BC063116 Homo sapi |
| 11         | 1284   | 56.9        | 1446   | 12 AX335726 | AX335726 Synthetic |
| 12         | 1277.6 | 56.6        | 1914   | 10 BC012675 | BC012675 Mus muscu |
| 13         | 1277.4 | 56.6        | 1881   | 6 AX961894  | AX961894 Sequence  |
| 14         | 1185   | 52.5        | 1222   | 6 BD260616  | BD260616 49 human  |
| 15         | 1178.2 | 52.2        | 1456   | 6 AR380154  | AR380154 Sequence  |
| 16         | 1118.8 | 49.6        | 1865   | 6 AX961898  | AX961898 Sequence  |
| 17         | 1097.4 | 48.6        | 1549   | 10 AF033566 | AF033566 Mus muscu |
| 18         | 1097   | 48.6        | 21234  | 6 AR232169  | AR232169 Sequence  |
| 19         | 1097   | 48.6        | 21234  | 6 AR406008  | AR406008 Sequence  |

|      |        |      |        |              |                    |
|------|--------|------|--------|--------------|--------------------|
| 20   | 1097   | 48.6 | 21234  | 6 AR534244   | AR534244 Sequence  |
| 21   | 1097   | 48.6 | 21234  | 6 AX392829   | AX392829 Sequence  |
| c 22 | 1097   | 48.6 | 159266 | 2 AC026097   | AC026097 Homo sapi |
| c 23 | 1097   | 48.6 | 170624 | 9 AC022096   | AC022096 Homo sapi |
| 24   | 1009.6 | 44.8 | 1445   | 10 BC002220  | BC002220 Mus muscu |
| 25   | 995    | 44.1 | 1150   | 6 CQ718110   | CQ718110 Sequence  |
| 26   | 913.4  | 40.5 | 1414   | 10 AY653308  | AY653308 Mus muscu |
| 27   | 910    | 40.3 | 1416   | 10 BC079006  | BC079006 Rattus no |
| 28   | 880.8  | 39.0 | 1745   | 5 BX931976   | BX931976 Gallus ga |
| 29   | 824.4  | 36.5 | 1864   | 9 AB056421   | AB056421 Macaca fa |
| 30   | 814.8  | 36.1 | 1757   | 6 CQ715616   | CQ715616 Sequence  |
| 31   | 813.2  | 36.0 | 1773   | 9 BC031549   | BC031549 Homo sapi |
| 32   | 811.6  | 36.0 | 1834   | 9 HUMCLKIB   | L29219 Homo sapien |
| 33   | 768.8  | 34.1 | 1705   | 10 MMCLKMRNA | X57186 Mouse clk m |
| 34   | 765.6  | 33.9 | 1740   | 10 MUSSTYKIN | M38381 Mouse serin |
| 35   | 692    | 30.7 | 1946   | 5 CR48121    | CR48121 Xenopus t  |
| 36   | 677.8  | 30.0 | 1544   | 5 BX936001   | BX936001 Gallus ga |
| 37   | 669.4  | 29.7 | 1017   | 5 BX929601   | BX929601 Gallus ga |
| 38   | 644.4  | 28.6 | 1743   | 9 HUMCLKIA   | L29222 Homo sapien |
| 39   | 643.2  | 28.5 | 1750   | 9 HUMKINCD   | M59287 Human prote |
| 40   | 640.6  | 28.4 | 2254   | 6 AR531674   | AR531674 Sequence  |
| c 41 | 639.4  | 28.3 | 1535   | 11 BV179987  | BV179987 sqm10751  |
| 42   | 630.6  | 28.0 | 1643   | 10 MUSMCLK   | L29221 Mus musculu |
| 43   | 628.2  | 27.8 | 4862   | 10 BC062887  | BC062887 Mus muscu |
| 44   | 617    | 27.3 | 3056   | 10 BC081942  | BC081942 Rattus no |
| 45   | 595.2  | 26.4 | 1634   | 5 BX929386   | BX929386 Gallus ga |

ALIGNMENTS

RESULT 1  
AR232168  
LOCUS AR232168 2354 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 1 from patent US 6455291.  
ACCESSION AR232168  
VERSION AR232168.1 GI:27274056  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2354)  
AUTHORS Yan, C., Ye, J., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.  
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof  
JOURNAL Patent: US 6455291-A 1 24-SEP-2002;  
FEATURES  
source Location/Qualifiers  
1..2354  
/organism="unknown"  
/mol\_type="genomic DNA"

|                       |         |   |       |            |    |        |       |
|-----------------------|---------|---|-------|------------|----|--------|-------|
| Query Match           | 100.0%  | Score   | 2256; | DB         | 6; | Length | 2354; |
| Best Local Similarity | 100.0%; | Pred. No.   | 0;    | Mismatches | 0; | Indels | 0;    |
| Matches               | 2256;   | Conservative  | 0;    |            |    |        |       |
| Qy                    | 1       | GACACTCATCCAGTCATTATTTAGAACCAAGTCTTGAATGACGAGATTATCGGCAC    | 60    |            |    |        |       |
| Db                    | 72      | GACACTCATCCAGTCATTATTTAGAACCAAGTCTTGAATGACGAGATTATCGGCAC    | 131   |            |    |        |       |
| Qy                    | 61      | CGGAGATACGTTGACGAATACAGGAATGACTGTGTGAAGGATATGTTCTTAGACATTAT | 120   |            |    |        |       |
| Db                    | 132     | CGGAGATACGTTGACGAATACAGGAATGACTGTGTGAAGGATATGTTCTTAGACATTAT | 191   |            |    |        |       |
| Qy                    | 121     | CACAGACATTTGAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCGGACGAGG    | 180   |            |    |        |       |
| Db                    | 192     | CACAGACATTTGAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCGGACGAGG    | 251   |            |    |        |       |
| Qy                    | 181     | AGAAGCAGTCTCTAAAGGAAGCGCAATAGACACTGTTTCAAGTCATCAGTCAGTTCGAG | 240   |            |    |        |       |
| Db                    | 252     | AGAAGCAGTCTCTAAAGGAAGCGCAATAGACACTGTTTCAAGTCATCAGTCAGTTCGAG | 311   |            |    |        |       |
| Qy                    | 241     | AGCCACCGAAGAAAGATCCAGGATATAGAGGATGATGAGAGGGTTCACCTGATCTGT   | 300   |            |    |        |       |





REFERENCE 1 (bases 1 to 2354)  
AUTHORS Yan, C., Ye, J., Ketchum K.A., Di Francesco, V. and Beasley, E.M.  
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof  
JOURNAL Patent: US 6630337-A 1 07-OCT-2003;  
FEATURES Location/Qualifiers  
source 1..2354  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 2256; DB 6; Length 2354;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACTCATCCAGTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTTATCGGAC 60  
DB 72 GACACTCATCCAGTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTTATCGGAC 131  
QY 61 CGGAGATAGCTTGAACCAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT 120  
DB 132 CGGAGATAGCTTGAACCAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT 191  
QY 121 CACAGAGATTTGAAGCGGATCGAATCCACTGAGTAAATCTTCAGTCGCGACGAG 180  
DB 192 CACAGAGATTTGAAGCGGATCGAATCCACTGAGTAAATCTTCAGTCGCGACGAG 251  
QY 181 AGACAGCTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGGAAG 240  
DB 252 AGACAGCTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGGAAG 311  
QY 241 AGCCACGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT 300  
DB 312 AGCCACGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT 371  
QY 301 CAAGTGAAGAGCTTCTAAGAGCAAGATAGAAATCGTGGACACCTTGGGTGAAGAGCC 360  
DB 372 CAAGTGAAGAGCTTCTAAGAGCAAGATAGAAATCGTGGACACCTTGGGTGAAGAGCC 431  
QY 361 TTTGGCAAGTTGTAGAGTGCAATCATCATGCGATGGATGGATGCGATGAGCAGTGAAG 420  
DB 432 TTTGGCAAGTTGTAGAGTGCAATCATCATGCGATGGATGGATGCGATGAGCAGTGAAG 491  
QY 421 ATCGTAAAGAAATGTAGCGCTTACCGTGAAGCAGCTCGTTCAAGAAATCCAAATATTAG 480  
DB 492 ATCGTAAAGAAATGTAGCGCTTACCGTGAAGCAGCTCGTTCAAGAAATCCAAATATTAG 551  
QY 481 CACTTAAATAGTACTGATCCCAATAGTCTTCCGATGTGTCAGATGTCAGATGAGTTT 540  
DB 552 CACTTAAATAGTACTGATCCCAATAGTCTTCCGATGTGTCAGATGTCAGATGAGTTT 611  
QY 541 GATCATCATGGTCATGTTGTATGTTGTTGAATCTGAGCTAGTGGACTTACGATTTTC 600  
DB 612 GATCATCATGGTCATGTTGTATGTTGTTGAATCTGAGCTAGTGGACTTACGATTTTC 671  
QY 601 ATTAAGAAAAACAGCTTTCTGCATTTCAAATTTGACCAATGACATCAGGAGATGCGTATCAG 660  
DB 672 ATTAAGAAAAACAGCTTTCTGCCATTTCAAATTTGACCAATGACATCAGGAGATGCGTATCAG 731  
QY 661 ATCTGCCAGTCAATAAATTTTTCATCATATAAATTTAAATCCCATACAGATCTGAAGCCT 720  
DB 732 ATCTGCCAGTCAATAAATTTTTCATCATATAAATTTAAATCCCATACAGATCTGAAGCCT 791  
QY 721 GAAATATTTTGTGTAAGTCTGACATGATGATGATGATGATGATGATGATGATGATGAT 780  
DB 792 GAAATATTTTGTGTAAGTCTGACATGATGATGATGATGATGATGATGATGATGATGAT 851  
QY 781 GATGAACGCACACTGAAAAACACAGATATCAAGATTTGTTGACTTTGGAGTGCACATGAT 840  
DB 852 GATGAACGCACACTGAAAAACACAGATATCAAGATTTGTTGACTTTGGAGTGCACATGAT 911  
QY 841 GATGATGAACATCAGTACTTTGTGTCTACCCGCGCATCAGAGCTCCCGAGTCAAT 900

DB 912 GATGATGAACATCAGTACTTTTGGTGTCTACCCGCGCATCAGAGCTCCCGAGTCAAT 971  
QY 901 TTGGCTTTAGGTGTGCTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTGAA 960  
DB 972 TTGGCTTTAGGTGTGCTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTGAA 1031  
QY 961 TATTACCTTTGGTTTTCAGAGTCTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATG 1020  
DB 1032 TATTACCTTTGGTTTTCAGAGTCTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATG 1091  
QY 1021 GAAACGAATATTAGGACCCATACCAACACACATGATTTCAGAAAAACAAGAAACGCAAGTAT 1080  
DB 1092 GAAACGAATATTAGGACCCATACCAACACACATGATTTCAGAAAAACAAGAAACGCAAGTAT 1151  
QY 1081 TTTTCAACCAATACAGCTAGATTGGGATGAACACACAGTCTCTGCTGGTGTAGATATGTTAGGAGA 1140  
DB 1152 TTTTCAACCAATACAGCTAGATTGGGATGAACACACAGTCTCTGCTGGTGTAGATATGTTAGGAGA 1211  
QY 1141 CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCTCATGATGAAGAAACATGAGAAACCTGTTT 1200  
DB 1212 CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCTCATGATGAAGAAACATGAGAAACCTGTTT 1271  
QY 1201 GACCTGGTTCCGAAGATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCA 1260  
DB 1272 GACCTGGTTCCGAAGATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCA 1331  
QY 1261 TTGAGCATCTCTTTTGAATTTTAAAGAAATGAAATGGAATGGAATGGAATGGAATGGAAT 1320  
DB 1332 TTGAGCATCTCTTTTGAATTTTAAAGAAATGAAATGGAATGGAATGGAATGGAATGGAAT 1391  
QY 1321 TATATATCTCTTGAAGAGATTTCTTAAAGCTGTGTGTCAGTCAACTAAACATCTTAAATAT 1380  
DB 1392 TATATATCTCTTGAAGAGATTTCTTAAAGCTGTGTGTCAGTCAACTAAACATCTTAAATAT 1451  
QY 1381 TTTTGTAAACATTAATTTATTTTGTACAGTTAAAGTAAATATTTGTATGTTTGTATCAAA 1440  
DB 1452 TTTTGTAAACATTAATTTATTTTGTACAGTTAAAGTAAATATTTGTATGTTTGTATCAAA 1511  
QY 1441 TAGCATATTAATCTTTTAAAGCAAGTATGCTTGTATGATGATGATGATGATGATGATGAT 1500  
DB 1512 TAGCATATTAATCTTTTAAAGCAAGTATGCTTGTATGATGATGATGATGATGATGATGAT 1571  
QY 1501 AATTTTCTTTTGAATTAACATTTTAAATATCTTTTGAATATCTTTTGTGTGTCAGTG 1560  
DB 1572 AATTTTCTTTTGAATTAACATTTTAAATATCTTTTGAATATCTTTTGTGTGTCAGTG 1631  
QY 1561 ATAAATGTGATGATCTTGGCTTTGTACAGGAGTCACTCTGAAAGTATTTTGTGTTT 1620  
DB 1632 ATAAATGTGATGATCTTGGCTTTGTACAGGAGTCACTCTGAAAGTATTTTGTGTTT 1691  
QY 1621 AGTAAAGGAAATCTTGACTACTTTTATATTTCTTAAAGGAAATTTCTTTATATATCTTCAA 1680  
DB 1692 AGTAAAGGAAATCTTGACTACTTTTATATTTCTTAAAGGAAATTTCTTTATATATCTTCAA 1751  
QY 1681 TTTAGAACTTAATCTTTTAAAGTCTTCTGTGTAATTTGTAACCGGTGATTTATTATTA 1740  
DB 1752 TTTAGAACTTAATCTTTTAAAGTCTTCTGTGTAATTTGTAACCGGTGATTTATTATTA 1811  
QY 1741 CTCTAGATAAGCAGTACTAGAACCAACCAAACTCTAGAAATCTTTTACTCTTGTAGAAATCTAT 1800  
DB 1812 CTCTAGATAAGCAGTACTAGAACCAACCAAACTCTAGAAATCTTTTACTCTTGTAGAAATCTAT 1871  
QY 1801 TAAATTTTAAAGTGTGTTATCTTTTCAATGGGTGATGTCAGGGTGTATAACACAGACATTC 1860  
DB 1872 TAAATTTTAAAGTGTGTTATCTTTTCAATGGGTGATGTCAGGGTGTATAACACAGACATTC 1931  
QY 1861 ATGAAAGGCATGAGTTTGTCCATTTGTGACAGTTTGTGTTTAAATAAACAACATACACACT 1920  
DB 1932 ATGAAAGGCATGAGTTTGTCCATTTGTGACAGTTTGTGTTTAAATAAACAACATACACACT 1991  
QY 1921 TTAATTAAGATTAATACTAATCTAGAAAGTCACTTGGAAATGACATTTTCCAGTATG 1980  
DB 1992 TTAATTAAGATTAATACTAATCTAGAAAGTCACTTGGAAATGACATTTTCCAGTATG 2051

```
QY 1981 TTTGGTGAAGTCAAGATATAAAATAGAAATCTGATGAGAGGTTTCAGTTTTTAATACC 2040
Db 2052 TTTGGTGAAGTCAAGATATAAAATAGAAATCTGATGAGAGGTTTCAGTTTTTAATACC 2111
QY 2041 AAGTCTTTAGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAA 2100
Db 2112 AAGTCTTTAGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAA 2171
QY 2101 ACCTATAAGAAATTAAGTTTATTAATATAGGCAATTTATGCTGTGATTAATTTCTTACGGGAG 2160
Db 2172 ACCTATAAGAAATTAAGTTTATTAATATAGGCAATTTATGCTGTGATTAATTTCTTACGGGAG 2231
QY 2161 AAAGAGGATTTGATTTGGAAGAGCTTTGGGAAGAAAGTCTGTAATTTCCAGAAATTT 2220
Db 2232 AAAGAGGATTTGATTTGGAAGAGCTTTGGGAAGAAAGTCTGTAATTTCCAGAAATTT 2291
QY 2221 AATTGATTTGGTTTACATAAATCTTTTGAATTTTGCATTTTCAGAAA 2256
Db 2292 AATTGATTTGGTTTACATAAATCTTTTGAATTTTGCATTTTCAGAAA 2327

RESULT 3
AR534243
LOCUS AR534243 2354 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6733978.
ACCESSION AR534243
VERSION AR534243.1 GI:53924333
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2354)
AUTHORS Yan, C., Ye, J., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding
human kinase proteins, and uses thereof
JOURNAL Patent: US 6733978-A 1 11-MAY-2004;
FEATURES
source Location/Qualifiers
1..2354
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 2256; DB 6; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACTCATCCCGAGTCATTTATTTAGAACGAAGGTCCTTGAATGAGCGAGATTTATCGGAC 60
Db 72 GACACTCATCCCGAGTCATTTATTTAGAACGAAGGTCCTTGAATGAGCGAGATTTATCGGAC 131
QY 61 CGGAGATCGTTGACGAATACAGGATGACTACTGTGAAGGATATGTTCTTAGACATTTAT 120
Db 132 CGGAGATCGTTGACGAATACAGGATGACTACTGTGAAGGATATGTTCTTAGACATTTAT 191
QY 121 CACAGAGACATTTGAACCGGGTATCGAATCCACTGCAAGTAAATCTTCAGTCCGACGAGG 180
Db 192 CACAGAGACATTTGAACCGGGTATCGAATCCACTGCAAGTAAATCTTCAGTCCGACGAGG 251
QY 181 AGAAGCAGTCTTAAAGAGGCGCAATAGACACTGTTTCAAGTCATCAGTCAGTTTCGAAG 240
Db 252 AGAAGCAGTCTTAAAGAGGCGCAATAGACACTGTTTCAAGTCATCAGTCAGTTTCGAAG 311
QY 241 AGCCACCGAGGAAGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGT 300
Db 312 AGCCACCGAGGAAGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGT 371
QY 301 CAAAGTGGAGACGCTTCTAAGAGCAAGATATGAATCGTGACACTTTGGGTGAAGGAGCC 360
Db 372 CAAAGTGGAGACGCTTCTAAGAGCAAGATATGAATCGTGACACTTTGGGTGAAGGAGCC 431
QY 361 TTGGGCAAGTTGTAGAGTGCAATTTGATCGGATCGGATGGCATGTAGCAGTGA 420
```

```
Db 432 TTTGGCAAAAGTTGTAGAGTGCATTTGATCGCATGGATGGCATGCGATGAGCAGTGAAA 491
QY 421 ATCGTAAAAAATCTAGGCCGTTTACCGTGAAGCAGCTCGTTTCAGAAAATCCAAAGTATTAGAG 480
Db 492 ATCGTAAAAAATCTAGGCCGTTTACCGTGAAGCAGCTCGTTTCAGAAAATCCAAAGTATTAGAG 551
QY 481 CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGTCAGATGGTTTT 540
Db 552 CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGTCAGATGGTTTT 611
QY 541 GATCATCATGTCATGTTTGTATTTGTGTTTGAAGTCTGAGTCTGAGTCTTACGATTTTC 600
Db 612 GATCATCATGTCATGTTTGTATTTGTGTTTGAAGTCTGAGTCTTACGATTTTC 671
QY 601 ATTAAAGAAAACAGCTTTCTGCCAATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAG 660
Db 672 ATTAAAGAAAACAGCTTTCTGCCAATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAG 731
QY 661 ATCTGCCAGTCAATAAATTTTTTATCATATAAATAAATTAACCCATACAGATCTGAAGCCT 720
Db 732 ATCTGCCAGTCAATAAATTTTTTATCATATAAATAAATTAACCCATACAGATCTGAAGCCT 791
QY 721 GAAAATATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAATGAAACGT 780
Db 792 GAAAATATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAATGAAACGT 851
QY 781 GATGAACGCACACTGAAAAAACAAGATATCAAAAGTTTGTGACCTTTGGAAGTGAACAGTAT 840
Db 852 GATGAACGCACACTGAAAAAACAAGATATCAAAAGTTTGTGACCTTTGGAAGTGAACAGTAT 911
QY 841 GATGATGAAACATACAGTACTTTTGTGTCTACCCGGCACTACAGAGCTCCGAGGTCAAT 900
Db 912 GATGATGAAACATACAGTACTTTTGTGTCTACCCGGCACTACAGAGCTCCGAGGTCAAT 971
QY 901 TTGCTCTTTAGTGTGCTCTCAGCCTTTGTGAGTGTGAGCAGTAGTGTGCATTTCTTATTGAA 960
Db 972 TTGCTCTTTAGTGTGCTCTCAGCCTTTGTGAGTGTGAGCAGTAGTGTGCATTTCTTATTGAA 1031
QY 961 TATTACCTTTGGTTTTCAGAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGGCAATGATG 1020
Db 1032 TATTACCTTTGGTTTTCAGAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGGCAATGATG 1091
QY 1021 GAACGAATATTAGGACCCATACCAACACATGATTCAGAAAAACAAGAAAAACGCAAGTAT 1080
Db 1092 GAACGAATATTAGGACCCATACCAACACATGATTCAGAAAAACAAGAAAAACGCAAGTAT 1151
QY 1081 TTTCAACCATTAACAGCTAGATTGGGATGAACACAGTTCTGCTGTGTAGATATGTTTAGGAGA 1140
Db 1152 TTTCAACCATTAACAGCTAGATTGGGATGAACACAGTTCTGCTGTGTAGATATGTTTAGGAGA 1211
QY 1141 CGCTGCAAAACCGTTGAAGGAATTTTATGCTTTGTCTATGATGAAGAACATGAGAAAATGTTT 1200
Db 1212 CGCTGCAAAACCGTTGAAGGAATTTTATGCTTTGTCTATGATGAAGAACATGAGAAAATGTTT 1271
QY 1201 GACCTGGTTTCAAGAAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTCGGATGAAGCA 1260
Db 1272 GACCTGGTTTCAAGAAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTCGGATGAAGCA 1331
QY 1261 TTGAGCATCTCTTTCTTTGACTTTATTTAAAAAAGAAATGAAAATGGGAATTCAGTGGTCTTAC 1320
Db 1332 TTGAGCATCTCTTTCTTTGACTTTATTTAAAAAAGAAATGAAAATGGGAATTCAGTGGTCTTAC 1391
QY 1321 TATATACCTCTCTAGAGAGATTTTAAAGCTGTGTGAGTAAATTTGATGTTTGTATCAAA 1380
Db 1392 TATATACCTCTCTAGAGAGATTTTAAAGCTGTGTGAGTAAATTTGATGTTTGTATCAAA 1451
QY 1381 TTTTGTGAACATTAATTTATTTTGTACAGTTAAAGTAAATTTGATGTTTGTATCAAA 1440
Db 1452 TTTTGTGAACATTAATTTATTTTGTACAGTTAAAGTAAATTTGATGTTTGTATCAAA 1511
QY 1441 TAGCATAAATTAACCTTTGTTTGAAGCAAGTATGGTCTTGTGATTAATGCAATTAAGAAAAAT 1500
Db 1512 TAGCATAAATTAACCTTTGTTTGAAGCAAGTATGGTCTTGTGATTAATGCAATTAAGAAAAAT 1571
```



961 TATTACCTTGGTTTTCACAGTCTTTTTCAGACTCATGATAGTAAGAGCACCTCGCAATGATG 1020  
 1032 TATTACCTTGGTTTTCACAGTCTTTTTCAGACTCATGATAGTAAGAGCACCTCGCAATGATG 1091  
 1021 GAACGAATATTAGAGCCATACCAACACATGATTCAGAAAAACAGAAACGCAAGTAT 1080  
 1092 GAACGAATATTAGAGCCATACCAACACATGATTCAGAAAAACAGAAACGCAAGTAT 1151  
 1081 TTTTCCACATACCAAGTATGGATGAAACACAGTCTTCTGCTGGTATGATATGTTAGGAGA 1140  
 1152 TTTTCCACATACCAAGTATGGATGAAACACAGTCTTCTGCTGGTATGATATGTTAGGAGA 1211  
 1141 CGCTGCAACCGTTTGAAGAAATTTATGCTTTTTCATGATAGAACATGAGAAACGTTT 1200  
 1212 CGCTGCAACCGTTTGAAGAAATTTATGCTTTTTCATGATAGAACATGAGAAACGTTT 1271  
 1201 GACCTGGTTTGAAGAAATTTTATGATATGATCAACTCAAGAAATTAACCTTGGATGAAGCA 1260  
 1272 GACCTGGTTTGAAGAAATTTTATGATATGATCAACTCAAGAAATTAACCTTGGATGAAGCA 1331  
 1261 TTGAGAGCATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320  
 1332 TTGAGAGCATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1391  
 1321 TATATACCTTCTTGAAGAGATTTACTTAAAGACTGTGTCACTCAACTCAACTCAACTCAACT 1380  
 1392 TATATACCTTCTTGAAGAGATTTACTTAAAGACTGTGTCACTCAACTCAACTCAACTCAACT 1451  
 1381 TTTTGTAAACATTAATTTTGTATAGTTAAAGTAAATTTTGTATGTTTGTATGATCAAA 1440  
 1452 TTTTGTAAACATTAATTTTGTATAGTTAAAGTAAATTTTGTATGTTTGTATGATCAAA 1511  
 1441 TAGCATATTAACCTTGTAAAGCAAGTATGTTTGTATGATGATGATGATGATGATGATGAT 1500  
 1512 TAGCATATTAACCTTGTAAAGCAAGTATGTTTGTATGATGATGATGATGATGATGATGAT 1571  
 1501 AATTTTCTTTTGAATATACCATTTTAAATACCTTTTAAATATCCTTTTGAATATCCTTTT 1560  
 1572 AATTTTCTTTTGAATATACCATTTTAAATATCCTTTTGAATATCCTTTTGAATATCCTTTT 1631  
 1561 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 1632 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1691  
 1621 AGTAAAGCAATCTTGTACTTCTTATATCTTAAAGCAATCTTGTATATATCTTGTATAT 1680  
 1692 AGTAAAGCAATCTTGTACTTCTTATATCTTAAAGCAATCTTGTATATATCTTGTATAT 1751  
 1681 TTTAGAACTTAACTTTAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1740  
 1752 TTTAGAACTTAACTTTAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1811  
 1741 CTCTAGATAGCAGGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGATTTCTAT 1800  
 1812 CTCTAGATAGCAGGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGATTTCTAT 1871  
 1801 TAAATTTTAAAGTGTGTTATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860  
 1872 TAAATTTTAAAGTGTGTTATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1931  
 1861 ATGAAAGGCATGCAAGTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1920  
 1932 ATGAAAGGCATGCAAGTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1991  
 1921 TTAATTTAAGATTAATCTAATCTAATCTAATCTAATCTAATCTAATCTAATCTAATCTAAT 1980  
 1992 TTAATTTAAGATTAATCTAATCTAATCTAATCTAATCTAATCTAATCTAATCTAATCTAAT 2051  
 1981 TTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
 2052 TTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2111  
 2041 AAGTCTTTAGGAGTCTTAACTTTGAGGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2100

2112 AAGTCTTTAGGAGTCTTAACTATGCGCAGCATCTGTTTATCAAAATGACATAAATACGTAA 2171  
 2101 ACCTATAGAAATTAAGTTTATTAATTAAGCAATTTATGCTGTGATATTTCTTACGGGAG 2160  
 2172 ACCTATAGAAATTAAGTTTATTAATTAAGCAATTTATGCTGTGATATTTCTTACGGGAG 2231  
 2161 AAAGAGGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTT 2220  
 2232 AAAGAGGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTT 2291  
 2221 AATTGATTTGTTTACATAACTTTTGTGACTTCAGAAA 2256  
 2292 AATTGATTTGTTTACATAACTTTTGTGACTTCAGAAA 2327

RESULT 5  
 AX056410  
 LOCUS  
 DEFINITION Sequence 54 from Patent WO0073469.  
 ACCESSION AX056410  
 VERSION AX056410.1 GI:12229117  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 Plozman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.  
 Protein Kinases  
 TITLE  
 JOURNAL Patent: WO 0073469-A 54 07-DEC-2000;  
 Sugen, Inc. (US)  
 FEATURES  
 source  
 1. 2488  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 90.8%; Score 2048; DB 6; Length 2488;  
 Best local Similarity 95.9%; Pred. No. 0;  
 Matches 2148; Conservative 0; Mismatches 1; Indels 90; Gaps 1;

14 GTCATTATTAGAACCAAGTCTTGAATAGCAGATATTATCGGACCGGAGATACGTTG 73  
 338 GTCATTATTAGAACCAAGTCTTGAATAGCAGATATTATCGGACCGGAGATACGTTG 397  
 74 ACGATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGAGACATTG 133  
 398 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGAGACATTG 457  
 134 AAACGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGACGAGGAGAGAGCAGTCTTA 193  
 458 AAACGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGACGAGGAGAGAGCAGTCTTA 517  
 194 AAAGAAAGCGCAATAGACATCTTTCAAGTATCATGTCAGTTCGGAAGACCCGGAAGGA 253  
 518 AAAGAAAGCGCAATAGACATCTTTCAAGTATCATGTCAGTTCAGTTCAGTTCAGTTCAG 561  
 254 AAAGTCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGAG 313  
 562 ----- 561  
 314 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 373  
 562 -----NATGAATCGTGGACACTTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 607  
 374 TAGAGTGCATTGATCATGGCATGGATGCGATGCGATGAGTGAAGTGAAGTGAAGTGAAGTGA 433  
 608 TAGAGTGCATTGATCATGGCATGGATGCGATGCGATGAGTGAAGTGAAGTGAAGTGAAGTGA 667  
 434 TAGGCCGTACCGTGAAGCAGCTGTTTCAGAAATCCAGATATTAGAGCATTAAATAGTA 493

Db 668 TAGGCGGTACCGTGAAGAGCGCTCTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 727  
Qy 494 CTGATCCCAATAGTGTCTTCGATGTGCCAGATGCTAGAAATGGTTTGATCATCATGTC 553  
Db 728 CTGATCCCAATAGTGTCTTCGATGTGCCAGATGCTAGAAATGGTTTGATCATCATGTC 787  
Qy 554 ATGTTTGTATTGTGTTGAACTGACCTAGTGGGACCTAGTACTTACGATTTCAATTAAGAAACA 613  
Db 788 ATGTTTGTATTGTGTTGAACTGACCTAGTGGGACCTAGTACTTACGATTTCAATTAAGAAACA 847  
Qy 614 GCTTTCTGCACTTTCAAATTTGACACATCAGGACATGCGGTATCAGATCTGCCAGTCAA 673  
Db 848 GCTTTCTGCACTTTCAAATTTGACACATCAGGACATGCGGTATCAGATCTGCCAGTCAA 907  
Qy 674 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAAGCTGAAATATTTTGT 733  
Db 908 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAAGCTGAAATATTTTGT 967  
Qy 734 TTGTGAAGTCTGACTATGTAGTCAATATATATCTTAAATGAAAGCTGATGAACGCAC 793  
Db 968 TTGTGAAGTCTGACTATGTAGTCAATATATATCTTAAATGAAAGCTGATGAACGCAC 1027  
Qy 794 TGAATAACACAGATATCAAGTTGTGTGACTTTTGAAGTGCACGATGATGATGAACATC 853  
Db 1028 TGAATAACACAGATATCAAGTTGTGTGACTTTTGAAGTGCACGATGATGATGAACATC 1087  
Qy 854 ACAGTACTTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 913  
Db 1088 ACAGTACTTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 1147  
Qy 914 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTAATTAATTAACCTTGGTT 973  
Db 1148 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTAATTAATTAACCTTGGTT 1207  
Qy 974 TCACAGTCTTTCAGCTCATGTAGTAAAGACCTGCGCATGTGAAAGATTAATTAAG 1033  
Db 1208 TCACAGTCTTTCAGCTCATGTAGTAAAGACCTGCGCATGTGAAAGATTAATTAAG 1267  
Qy 1034 GACCATACCAACACACATGATTTCAGAAAAACGAAAGTATTTTCAACATAACC 1093  
Db 1268 GACCATACCAACACACATGATTTCAGAAAAACGAAAGTATTTTCAACATAACC 1327  
Qy 1094 AGCTAGATTGGATGAAACACAGTCTGCTGTGTAGATATGTTAGGAGACGCTGCAACCGT 1153  
Db 1328 AGCTAGATTGGATGAAACACAGTCTGCTGTGTAGATATGTTAGGAGACGCTGCAACCGT 1387  
Qy 1154 TGAAGGAATTTATGCTTGTCTATGATGAGACATGAGAACTGTTGACCTGGTTCGAA 1213  
Db 1388 TGAAGGAATTTATGCTTGTCTATGATGAGACATGAGAACTGTTGACCTGGTTCGAA 1447  
Qy 1214 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGAGCATCCTT 1273  
Db 1448 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGAGCATCCTT 1507  
Qy 1274 TCTTTGACTTATTAATAAGAAATGAAATGGAATCAGTGGTCTTATATATATCTCTCT 1333  
Db 1508 TCTTTGACTTATTAATAAGAAATGAAATGGAATCAGTGGTCTTATATATATCTCTCTCT 1567  
Qy 1334 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATCTAATATTTTGTAAACAT 1393  
Db 1568 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATCTAATATTTTGTAAACAT 1627  
Qy 1394 AAATTTATTTGTACAGTTAAAGTAAATATGATGTTTGTATCAATAGCATAATTAAC 1453  
Db 1628 AAATTTATTTGTACAGTTAAAGTAAATATGATGTTTGTATCAATAGCATAATTAAC 1687  
Qy 1454 TTGTTAAGCAAGTATGTTGATTAATGCAATAGAAAAATTAATAATTTTCTTTT 1513  
Db 1688 TTGTTAAGCAAGTATGTTGATTAATGCAATAGAAAAATTAATAATTTTCTTTT 1747  
Qy 1514 GAAATACCAATTTTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATATGATG 1573  
Db 1748 GAAATACCAATTTTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATATGATG 1807

Qy 1574 ATCTTGCCCTTTGTACATGGAGGTCACTCTGAAAGTGAATTTTGTGAGTAAAGGAAT 1633  
Db 1808 ATCTTGCCCTTTGTACATGGAGGTCACTCTGAAAGTGAATTTTGTGAGTAAAGGAAT 1867  
Qy 1634 CTTGACTACTTTATATCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAAC 1693  
Db 1868 CTTGACTACTTTATATCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAAC 1927  
Qy 1694 TTTAAAGTCTTTCTCTGTAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCA 1753  
Db 1928 TTTAAAGTCTTTCTCTGTAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCA 1987  
Qy 1754 GGTACTAGAAACCAAACTCAGAAAAATGTTTCTGTTAGAAATCTTATTTAAATTTTAAAGTG 1813  
Db 1988 GGTACTAGAAACCAAACTCAGAAAAATGTTTCTGTTAGAAATCTTATTTAAATTTTAAAGTG 2047  
Qy 1814 TTGTAATCTTTTCTTATTTGGGTGATGTCAGGGTGATTAACGACATTCATGGAAGGCATG 1873  
Db 2048 TTGTAATCTTTTCTTATTTGGGTGATGTCAGGGTGATTAACGACATTCATGGAAGGCATG 2107  
Qy 1874 CAGTTTGTCCATCTGACAGTTTGTAAATAAAACACATACACACTTTTATTTAAAGATTA 1933  
Db 2108 CAGTTTGTCCATCTGACAGTTTGTAAATAAAACACATACACACTTTTATTTAAAGATTA 2167  
Qy 1934 AAATCTAACTGGAAGTCACTTGGAAAAATGGAATTTTCCAAAGTATGTTTGGTGAATC 1993  
Db 2168 AAATCTAACTGGAAGTCACTTGGAAAAATGGAATTTTCCAAAGTATGTTTGGTGAATC 2227  
Qy 1994 AGATATAAAATAGAAATCTGATGAGAGGTTTCAGTTTTTAAATACCAAGTCCCTTAGGAG 2053  
Db 2228 AGATATAAAATAGAAATCTGATGAGAGGTTTCAGTTTTTAAATACCAAGTCCCTTAGGAG 2287  
Qy 2054 TCTTAAACATTCGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAAT 2113  
Db 2288 TCTTAAACATTCGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAAT 2347  
Qy 2114 AAGTTTATTAATTAGGCAATTTATGTCGTGATAATCTTACGGGAGAAAGGATTTGA 2173  
Db 2348 AAGTTTATTAATTAGGCAATTTATGTCGTGATAATCTTACGGGAGAAAGGATTTGA 2407  
Qy 2174 TTGGAAGCAGTTTGGGAGAAAGTCTGCTGTAATTTCCAGAAATTTCCAGAAATTTAATGATTTG 2233  
Db 2408 TTGGAAGCAGTTTGGGAGAAAGTCTGCTGTAATTTCCAGAAATTTAATGATTTG 2467  
Qy 2234 CATAACTTTTGTGACTTCA 2252  
Db 2468 CATAACTTTTGTGACTTCA 2486

## RESULT 6

AR077239 LOCUS AR077239 2446 bp DNA linear PAT 31-AUG-2000  
DEFINITION Sequence 9 from patent US 5962232.

AR077239 ACCESSION AR077239.1 GI:10003985

AR077239.1 VERSION AR077239.1

AR077239.1 KEYWORDS GI:10003985

AR077239.1 SOURCE Unknown.

AR077239.1 ORGANISM Unknown.

AR077239.1 UNCLASSIFIED.

AR077239.1 1 (bases 1 to 2446)

AR077239.1 AUTHORS Bandman, O., Hillman, J.L., Lal, P., Akerblom, I.E., Shah, P.,

AR077239.1 TITLE Corley, N.C. and Guegler, K.J.

AR077239.1 JOURNAL Protein kinase molecules

AR077239.1 PATENT: US 5962232-A 9 05-OCT-1999;

AR077239.1 FEATURES Location/Qualifiers

AR077239.1 1..2446

AR077239.1 /organism="unknown"

AR077239.1 /mol\_type="unassigned DNA"

## ORIGIN

Query Match 90.7%; Score 2046.6; DB 6; Length 2446;  
Best Local Similarity 95.8%; Pred. No. 0;

| Matches 2149; Conservative 0; Mismatches 4; Indels 90; Gaps 1; |      |  |      |
|--|------|--|------|
| QY   | 14   | GTCATTATTTAGAACCAAGGTCCTTGAATGACGAGATTATTCGGACCGGAGATACGTTG    | 73   |
| Db   | 288  |  |      |
| QY   | 74   | ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTCTAGACATTTATCACAGACATTTG  | 133  |
| Db   | 348  | ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTCTAGACATTTATCACAGACATTTG  | 407  |
| QY   | 134  | AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCGAGGAGACGATCCTTA    | 193  |
| Db   | 408  | AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCGAGGAGACGATCCTTA    | 467  |
| QY   | 194  | AAAGGAGCGCAATAGACACATGTTCAAGTCATCAGTCACGTTTCGAAGAGCCACCAGGA    | 253  |
| Db   | 468  | AAAGGAGCGCAATAGACACATGTTCAAGTCATCAGTCACGTTTCGAAGTCACTGATGA     | 515  |
| QY   | 254  | AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAAAAGTGGAGCG     | 313  |
| Db   | 516  | -----  | 515  |
| QY   | 314  | TTCTAAGAGCAAGATATGAATCTGTGGACACTTTGGGTGAAGGAGCCTTTGGCAAGTTG    | 373  |
| Db   | 516  | -----AATCCGTGGACACTTTGGGTGAAGGAGCCTTTGGCAAGTTG                 | 557  |
| QY   | 374  | TAGAGTGCATTGATCATGGCATGGATGCGATGTCAGTACGATGAAATCGTAAAAAATG     | 433  |
| Db   | 558  | TAGAGTGCATTGATCATGGCATGGATGCGATGTCAGTACGATGAAATCGTAAAAAATG     | 617  |
| QY   | 434  | TAGGCCGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA | 493  |
| Db   | 618  | TAGGCCGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA | 677  |
| QY   | 494  | CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGATGTTTGCATCATCTGTC        | 553  |
| Db   | 678  | CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGATGTTTGCATCATCTGTC        | 737  |
| QY   | 554  | ATGTTTGTATTGTGTTTCAACTACTTGGGACTTAGTACTTACGATTTCAATTAAGAAAACA  | 613  |
| Db   | 738  | ATGTTTGTATTGTGTTTCAACTACTTGGGACTTAGTACTTACGATTTCAATTAAGAAAACA  | 797  |
| QY   | 614  | GCTTTCGCCATTTCAAAATTGACCAATCAGGAGATGCGGTATCAGATCTGCCAGTCAA     | 673  |
| Db   | 798  | GCTTTCGCCATTTCAAAATTGACCAATCAGGAGATGCGGTATCAGATCTGCCAGTCAA     | 857  |
| QY   | 674  | TAAATTTTTTACATCAATAATAAATTAACCCATACAGATCTGAAGCCTGAAAATATTTTGT  | 733  |
| Db   | 858  | TAAATTTTTTACATCAATAATAAATTAACCCATACAGATCTGAAGCCTGAAAATATTTTGT  | 917  |
| QY   | 734  | TTGTGAAGTCTGACTATCTAGTCAAAATATAATTTCTAAAATGAAACGTGATGAACGCAC   | 793  |
| Db   | 918  | TTGTGAAGTCTGACTATCTAGTCAAAATATAATTTCTAAAATGAAACGTGATGAACGCAC   | 977  |
| QY   | 794  | TGAAAAACACAGATATCAAAAGTTGTGTACTTTGGAAAGTGCACACGATGATGATGAACATC | 853  |
| Db   | 978  | TGAAAAACACAGATATCAAAAGTTGTGTACTTTGGAAAGTGCACACGATGATGATGAACATC | 1037 |
| QY   | 854  | ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTAGGTT     | 913  |
| Db   | 1038 | ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTAGGTT     | 1097 |
| QY   | 914  | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCATTTCTATTGAATATTACCTTGGTT   | 973  |
| Db   | 1098 | GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCATTTCTATTGAATATTACCTTGGTT   | 1157 |
| QY   | 974  | TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTCGCAATGATGAAACGAATATTAG   | 1033 |
| Db   | 1158 | TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTCGCAATGATGAAACGAATATTAG   | 1217 |
| QY   | 1034 | GACCCATACCAACACATGATTTCAGAAAACAGAAAACGCAAGTATTTTTCACCAATACC    | 1093 |
| Db   | 1218 | GACCCATACCAACACATGATTTCAGAAAACAGAAAACGCAAGTATTTTTCACCAATACC    | 1277 |

|    |      |   |      |
|----|------|---|------|
| QY | 1094 | AGCTAGATTGGGATGAACACACAGTTCTCTGCTGTAGATATGTTAGGAGACGCTGCAAAACCGT    | 1153 |
| Db | 1278 | AGCTAGATTGGGATGAACACACAGTTCTCTGCTGTAGATATGTTAGGAGACGCTGCAAAACCGT    | 1337 |
| QY | 1154 | TGAAGGAATTTATGCTTTGTTCATGATGAAGAACATGATGAAGAACTGTTTGACCTGGTTCGAA    | 1213 |
| Db | 1338 | TGAAGGAATTTATGCTTTGTTCATGATGAAGAACATGATGAAGAACTGTTTGACCTGGTTCGAA    | 1397 |
| QY | 1214 | GAATGTTAGAAATATGATCCAACTCAAAAGAAATTTACCTTGGATGAAGCATTGCGAGCATCCTT   | 1273 |
| Db | 1398 | GAATGTTAGAAATATGATCCAACTCAAAAGAAATTTACCTTGGATGAAGCATTGCGAGCATCCTT   | 1457 |
| QY | 1274 | TCTTTGACTTATTAAAAAGAAATGAAATGGAATCGAGTGGTCTTACTATATCTTCTCT          | 1333 |
| Db | 1458 | TCTTTGACTTATTAAAAAGAAATGAAATGGAATCGAGTGGTCTTACTATATCTTCTCT          | 1517 |
| QY | 1334 | AGAAGAGATTAATTAAGACATGTCAGTCAACATAAACTTAATTAATTTTGTAAACATTT         | 1393 |
| Db | 1518 | AGAAGAGATTAATTAAGACATGTCAGTCAACATAAACTTAATTAATTTTGTAAACATTT         | 1577 |
| QY | 1394 | AAATTTATTTGTACAGTTAAGTGAATATTTGTATGTTTCTTGTATCAATAGCATATTAAC        | 1453 |
| Db | 1578 | AAATTTATTTGTACAGTTAAGTGAATATTTGTATGTTTCTTGTATCAATAGCATATTAAC        | 1637 |
| QY | 1454 | TTGTTAAGCAAGTATGGTCTTTGATAATGCAATTAGAAAAAATTAATTTTCTTTT             | 1513 |
| Db | 1638 | TTGTTAAGCAAGTATGGTCTTTGATAATGCAATTAGAAAAAATTAATTTTCTTTT             | 1697 |
| QY | 1514 | GAATTTACCATTTTTAATACCTTTGAAATATCTTTGTGTCCAGTGATTAATGTGATTG          | 1573 |
| Db | 1698 | GAATTTACCATTTTTAATACCTTTGAAATATCTTTGTGTCCAGTGATTAATGTGATTG          | 1757 |
| QY | 1574 | ATCTTGCTTTTGTACATGAGGTCACTCTGAAAGTGAATTTTGTGAGTAAAAAGGAAAT          | 1633 |
| Db | 1758 | ATCTTGCTTTTGTACATGAGGTCACTCTGAAAGTGAATTTTGTGAGTAAAAAGGAAAT          | 1817 |
| QY | 1634 | CTTGACTACTTTTATTAATTTCTTAAAGGAATATTTCTTTATATACCTTCAAAATTTAGAACTTAAC | 1693 |
| Db | 1818 | CTTGACTACTTTTATTAATTTCTTAAAGGAATATTTCTTTATATACCTTCAAAATTTAGAACTTAAC | 1877 |
| QY | 1694 | TTTAAAAAGTTTCTCTCTGTAATTTGTTGAACGGGTGATTAATTAATCTCTAGATAAGCA        | 1753 |
| Db | 1878 | TTTAAAAAGTTTCTCTCTGTAATTTGTTGAACGGGTGATTAATTAATCTCTAGATAAGCA        | 1937 |
| QY | 1754 | GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAATTTCTATTAATTTTAAAGTG       | 1813 |
| Db | 1938 | GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAATTTCTATTAATTTTAAAGTG       | 1997 |
| QY | 1814 | TTGTATTCTTTTTCATTGGGTGATGTGAGGTGATAACCCAGACATTCATGGAAGGCATG         | 1873 |
| Db | 1998 | TTGTATTCTTTTTCATTGGGTGATGTGAGGTGATAACCCAGACATTCATGGAAGGCATG         | 2057 |
| QY | 1874 | CAGTTTGTCCATTTGTGACAGTTTGTATTAATAAAACCATACACACATTTTATTAAGATTA       | 1933 |
| Db | 2058 | CAGTTTGTCCATTTGTGACAGTTTGTATTAATAAAACCATACACACATTTTATTAAGATTA       | 2117 |
| QY | 1934 | AAATCTAACGGAAGTCAGCTTCGAAAAATGCAATTTCCAAATGATGTTTGTGAGTCAC          | 1993 |
| Db | 2118 | AAATCTAACGGAAGTCAGCTTCGAAAAATGCAATTTCCAAATGATGTTTGTGAGTCAC          | 2177 |
| QY | 1994 | AGATATAAAAAATAGAAATTTCTGATGAGAGTTTTCAGTTTAAATCCAAAGTCCCTTAGGAG      | 2053 |
| Db | 2178 | AGATATAAAAAATAGAAATTTCTGATGAGAGTTTTCAGTTTAAATCCAAAGTCCCTTAGGAG      | 2237 |
| QY | 2054 | TCTTAACTATGGCCAGCATCTGTTTATCAATGATCAATAAATACCTTAACCTTATAGAAAT       | 2113 |
| Db | 2238 | TCTTAACTATGGCCAGCATCTGTTTATCAATGATCAATAAATACCTTAACCTTATAGAAAT       | 2297 |
| QY | 2114 | AAGTTTATTAATTAAGGCAATTTATGTCGTGATAATTTCTTACGGGAGAAAGAGGATTTGA       | 2173 |
| Db | 2298 | AAGTTTATTAATTAAGGCAATTTATGTCGTGATAATTTCTTACGGGAGAAAGAGGATTTGA       | 2357 |









|    |      |  |      |
|----|------|--|------|
| Qy | 494  | CTGATCCCAATPAGTGTCTTCGATGTGTCGAGATGCTAGAAATGGTTTGGATCATCATGGCT     | 553  |
| Db | 641  | CTGATCCCAATPAGTGTCTTCGATGTGTCGAGATGCTAGAAATGGTTTGGATCATCATGGCT     | 700  |
| Qy | 554  | ATGTTTCTGATTGTCTTTGAACTACTCGGGACTTACTAGTCTTACGATTTTCATTTAAAGAAAACA | 613  |
| Db | 701  | ATGTTTGTATTGTGTTTGAACACTCTGGACCTTAGTACTTACGATTTTCATTTAAAGAAAACA    | 760  |
| Qy | 614  | GCTTTCTGCCAATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA     | 673  |
| Db | 761  | GCTTTCTGCCAATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA     | 820  |
| Qy | 674  | TAAATTTTTTACATCAATAATAAATTAACCCATACAGATCTGAAGCCTGAAAATAATTTTGT     | 733  |
| Db | 821  | TAAATTTTTTACATCAATAATAAATTAACCCATACAGATCTGAAGCCTGAAAATAATTTTGT     | 880  |
| Qy | 734  | TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAATGAAACGATGTAACGCACAC     | 793  |
| Db | 881  | TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAATGAAACGATGTAACGCACAC     | 940  |
| Qy | 794  | TGAAAAACACAGATATCAAAAGTTGTTGACTTTTGGAAGTGCAACGATGATGTAACGCACATC    | 853  |
| Db | 941  | TGAAAAACACAGATATCAAAAGTTGTTGACTTTTGGAAGTGCAACGATGATGTAACGCACATC    | 1000 |
| Qy | 854  | ACAGTACTTTGGTGTCTACCCGGCACATACAGAGTCCCGAGGTCATTTTGGCTTTTAGGTT      | 913  |
| Db | 1001 | ACAGTACTTTGGTGTCTACCCGGCACATACAGAGTCCCGAGGTCATTTTGGCTTTTAGGTT      | 1060 |
| Qy | 914  | GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGCATTTCTATTGGAATATTACCTTGGTT      | 973  |
| Db | 1061 | GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGCATTTCTATTGGAATATTACCTTGGTT      | 1120 |
| Qy | 974  | TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTTGGCAATGATGGAACGGAATATTAG     | 1033 |
| Db | 1121 | TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTTGGCAATGATGGAACGGAATATTAG     | 1180 |
| Qy | 1034 | GACCCATACCAACAACAATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACC     | 1093 |
| Db | 1181 | GACCCATACCAACAACAATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACC     | 1240 |
| Qy | 1094 | AGCTAGATTGGGATGAACACAGTCTCTCGTAGATATGTTAGGAGACGCTGCAAAACCGT        | 1153 |
| Db | 1241 | AGCTAGATTGGGATGAACACAGTCTCTCGTAGATATGTTAGGAGACGCTGCAAAACCGT        | 1300 |
| Qy | 1154 | TGAAGGAATTTATGCTTTTGTCATGATGAAGAAACAATGAGAAACTGTTTGAACCTGGTTCCAA   | 1213 |
| Db | 1301 | TGAAGGAATTTATGCTTTTGTCATGATGAAGAAACAATGAGAAACTGTTTGAACCTGGTTCCAA   | 1360 |
| Qy | 1214 | GAATGTTAGGAATATGATCTCAACCAAGAATTTACTCTTGGATGAAGCAATGCGAGATCCCT     | 1273 |
| Db | 1361 | GAATGTTAGGAATATGATCTCAACCAACGAAATTTACTCTTGGATGAAGCAATGCGAGATCCCT   | 1420 |
| Qy | 1274 | TCCTTGACTTTTAAAGAAATGAATGGGAATCAGTGGTCTTACTATATACTTCTCT            | 1333 |
| Db | 1421 | TCCTTGACTTTTAAAGAAATGAATGGGAATCAGTGGTCTTACTATATACTTCTCT            | 1480 |
| Qy | 1334 | AGAAGAGATTACTTTAAGACTGTGTCACTCAACTAAACAATTTCTAATATTTTGTGAAACATT    | 1393 |
| Db | 1481 | AGAAGAGATTACTTTAAGACTGTGTCACTCAACTAAACAATTTCTAATATTTTGTGAAACATT    | 1540 |
| Qy | 1394 | AAATTTATTTTGTACAGTTAAGTGAATAATGATATGTTTGTGTCAGTGCATAAATGTGATTG     | 1453 |
| Db | 1541 | AAATTTATTTTGTACAGTTAAGTGAATAATGATATGTTTGTGTCAGTGCATAAATGTGATTG     | 1600 |
| Qy | 1454 | TTGTTAGCAAGTATGTCCTTGATAATGCAATAGAAAAATTAATAATTTTCTTTTT            | 1513 |
| Db | 1601 | TTGTTAGCAAGTATGTCCTTGATAATGCAATAGAAAAATTAATAATTTTCTTTTT            | 1660 |
| Qy | 1514 | GAAATTAACCATTTTTTAATACCTTTTGAAATATATCCTTTGTGTCAGTGCATAAATGTGATTG   | 1573 |
| Db | 1661 | GAAATTAACCATTTTTTAATACCTTTTGAAATATATCCTTTGTGTCAGTGCATAAATGTGATTG   | 1720 |
| Qy | 1574 | ATCTTGCCCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTTTTTTTGTGATGAAAGGAAAT   | 1633 |

Db 1721 ATCTTGCCTTTTGTACATGAGGTCACCTCTGAAGTGATTTTTTTTGTAGTAAAGGAAT 1780  
 Qy 1634 CTTGACTACTTTATATCTTTAAAGGAATA 1662  
 Db 1781 CTTGACTACTTTTAAAAAAGAAAAA 1809

RESULT 9  
 AX961896  
 LOCUS  
 DEFINITION Sequence 107 from Patent WO03104277.  
 ACCESSION AX961896  
 VERSION AX961896.1 GI:40881306  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.  
 Stat6 activation gene  
 TITLE Patent: WO 03104277-A 107 18-DEC-2003;  
 JOURNAL Asahi Kasei Kabushiki Kaisha (JP)  
 FEATURES  
 source Location/Qualifiers  
 1..1792  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 1..1446  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAF06457.1"  
 /db\_xref="GI:40881307"  
 /translations="MRHSKRTCPDWDSRESWGSHYRGSHKRRSHSHSTQENRHKH  
 PHQFESDCHYLEARSLNDRYDRRYDYDEYKNDYCEGVVPRYHRDISSGYRIHCS  
 KSSVRSRRSPKRRHCHSCSHOSRSHKRKRSIEDDEGHLLICOSGDVLRYRVE  
 IVDLTGEGAFKQVECLDHGMDSHMVAVKLVNRYREARSEIQVLEHLSNTDNS  
 VFCVQMLWFDDHGHGVCIVFELGLSTDFIKENSFLPQIDHIFQMAIQICOSINF  
 LHNKHTULDKPENILFVKSDYVKNYSKMRDERTLNKTDIKVDFGSGATYDDDEHH  
 STLVNTHRYAPRPAIALGMSQCDVMSIGCILIEYLYLGTFTVFQTHDSKEHLLAMERI  
 LGP1PQMTQKTRKRYFHHNQLDWDDEHSSAGRYVRRRCKPLKEFMLCHDDEHEKLPD  
 LVRRMLEYDPTQRTITLDEALQHPFFDLLKKK"

CDS  
 ORIGIN

Query Match 72.2%; Score 1629.4; DB 6; Length 1792;  
 Best Local Similarity 99.9%; Pred. No. 3.9e-292;  
 Matches 1630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 14 GTCAATATTAGAACGAAGTCCTTGAATGAGCGAGATATTCGGACCGGAGATACGTTG 73  
 Db 161 GTCAATATTAGAACGAAGTCCTTGAATGAGCGAGATATTCGGACCGGAGATACGTTG 220

Qy 74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATATTCACAGAGACATTG 133  
 Db 221 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATATTCACAGAGACATTG 280

Qy 134 AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGAGAGACAGTCCTTA 193  
 Db 281 AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGAGAGACAGTCCTTA 340

Qy 194 AAAGGAAGCCCAATAGACACTGTTCAAGTCATCAGTCACGTTCCGAGAGCCACCGAAGGA 253  
 Db 341 AAAGGAAGCCCAATAGACACTGTTCAAGTCATCAGTCACGTTCCGAGAGCCACCGAAGGA 400

Qy 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACG 313  
 Db 401 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACG 460

Qy 314 TTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCGCTTTGGCAAGTTG 373  
 Db 461 TTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCGCTTTGGCAAGTTG 520



Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 51 Row: a Column: 9  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 10190705.

#### FEATURES

Location/Qualifiers  
1..3492  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30324121"  
/tissue\_type="skin and meninges pool- skin, dura matter,  
pia matter, choroid plexus."  
/clone\_lib="NIH MGC\_186"  
/lab\_host="DH10B"  
/note="Vector: pDNR-LIB"

#### ORIGIN

Query Match 58.4%; Score 1318.4; DB 9; Length 3492;  
Best Local Similarity 99.2%; Pred. No. 1.6e-234;  
Matches 1325; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 327 ATATGAATCGTGGACACTTTGGTGAAGAGCGCTTTGGCAAAGTTGTAGAGTGCATTGA 386  
Db AGATGAATCGTGGACACTTTGGTGAAGAGCGCTTTGGCAAAGTTGTAGAGTGCATTGA 2208  
QY 387 TCATGGCATGGATGGCATGCATGTAGCACTGAATCGTAAAAATGTAGGCGGTACCG 446  
Db TCATGGCATGGATGGCATGCATGTAGCACTGAATCGTAAAAATGTAGGCGGTACCG 2268  
QY 447 TGAAGCAGCTCGTTCAGAAATCAAGTATTAGAGCACTTAAATAGTACTGATCCCAATAG 506  
Db TGAAGCAGCTCGTTCAGAAATCAAGTATTAGAGCACTTAAATAGTACTGATCCCAATAG 2328  
QY 507 TGTCTTCCGATGTCCAGATCGTAAATGGTTTGATCATCATGTGTCATGTTGTATTGT 566  
Db TGTCTTCCGATGTCCAGATCGTAAATGGTTTGATCATCATGTGTCATGTTGTATTGT 2388  
QY 567 GTTTGAACCTACTGGGACTTGTAGTACTTACGATTCATTAAAGAAACAGCTTTCTGCCATT 626  
Db GTTTGAACCTACTGGGACTTGTAGTACTTACGATTCATTAAAGAAACAGCTTTCTGCCATT 2448  
QY 627 TCAAAATTGACACATCAGCAGATGCGGTATCAGATCTGCCAGTCAATAAATTTTACA 686  
Db TCAAAATTGACACATCAGCAGATGCGGTATCAGATCTGCCAGTCAATAAATTTTACA 2508  
QY 687 TCATAATAAATTAACCCATACAGATCTGAAGCTGAAATATTTTGTGTGAAGTCTGA 746  
Db TCATAATAAATTAACCCATACAGATCTGAAGCTGAAATATTTTGTGTGAAGTCTGA 2568  
QY 747 CTATGTAGTCAAAATATAATTTAAATGAAACGTGATGAACCGACACTGAAACACAGA 806  
Db CTATGTAGTCAAAATATAATTTAAATGAAACGTGATGAACCGACACTGAAACACAGA 2628  
QY 807 TATCAAGTTGTGACTTTGGAGTGCACGCTATGATGATGAACATCAGTACTTTGGT 866  
Db TATCAAGTTGTGACTTTGGAGTGCACGCTATGATGATGAACATCAGTACTTTGGT 2688  
QY 867 GTCTACCCGGCTACTACAGACTCCGAGGTCAATTTGGCTTGGTCTCAGCCTTG 926  
Db GTCTACCCGGCTACTACAGACTCCGAGGTCAATTTGGCTTGGTCTCAGCCTTG 2748  
QY 927 TGATGTTGGACATAGGTTGCACTTTATGAATATTACCTTGGTTTACAGTCTTTCA 986  
Db TGATGTTGGACATAGGTTGCACTTTATGAATATTACCTTGGTTTACAGTCTTTCA 2808  
QY 987 GACTCATGATAGTAAGACCACTGGCATGATGAACGAATATTAGNCCCATACCACA 1046  
Db GACTCATGATAGTAAGACCACTGGCATGATGAACGAATATTAGNCCCATACCACA 2868

QY 1047 ACACATGATTTCAGAAAAACAGAAACGCAAGTATTTTCCACATACCACGACTAGATTGGA 1106  
Db ACACATGATTTCAGAAAAACAGAAACGCAAGTATTTTCCACATACCACGACTAGATTGGA 2928  
QY 1107 TGAACACAGTCTGCTGCTGTAGATATGTTAGGAGACGCTGCAACCGTTCAAGAAATTTAT 1166  
Db TGAACACAGTCTGCTGCTGTAGATATGTTAGGAGACGCTGCAACCGTTCAAGAAATTTAT 2988  
QY 1167 GCTTTTGTGATGATGAAGAACATGAGAAACTGTTTGACCTGGTTCGAAGAATGTTAGAATA 1226  
Db GCTTTTGTGATGATGAAGAACATGAGAAACTGTTTGACCTGGTTCGAAGAATGTTAGAATA 3048  
QY 1227 TGATCCAACCTCAAGAAATTAACCTCGATGAAGCACTGCTCTTTGACTTAT 1286  
Db TGATCCAACCTCAAGAAATTAACCTCGATGAAGCACTGCTCTTTGACTTAT 3108  
QY 1287 AAAAAAGAAATGAAATGGAATCAGTGGTCTTACTATATATCTCTCTAGAGAGATTACT 1346  
Db AAAAAAGAAATGAAATGGAATCAGTGGTCTTACTATATATCTCTCTAGAGAGATTACT 3168  
QY 1347 TAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAATTTTGT 1406  
Db TAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAATTTTGT 3228  
QY 1407 CAGTTAAGTGAATATTTGTTGTTGTTATCAATAGCATTAATTAATTTGTTAAGCAAGT 1466  
Db CAGTTAAGTGAATATTTGTTGTTGTTATCAATAGCATTAATTAATTTGTTAAGCAAGT 3288  
QY 1467 ATGGTCTTGTAATGCAATAGAAATTAATTAATTTTCTTTTGAATTTACCATTT 1526  
Db ATGGTCTTGTAATGCAATAGAAATTAATTAATTTTCTTTTGAATTTACCATTT 3348  
QY 1527 TTAATATACCTTTGAATATCCTTTGTCAGTGAATAATGTTGATTTGTCCTTTTG 1586  
Db TTAATATACCTTTGAATATCCTTTGTCAGTGAATAATGTTGATTTGTCCTTTTG 3408  
QY 1587 TACATGGAGGTGACCTCTGAAAGTATTTTGTAGTAAAGAAATCTTGACTCTTTA 1646  
Db TACATGGAGGTGACCTCTGAAAGTATTTTGTAGTAAAGAAATCTTGACTCTTTA 3468  
QY 1647 TATCTTTAAGGAATA 1662  
Db AAAAAAATA 3484  
RESULT 11  
AY335726 1446 bp mRNA linear SYN 15-OCT-2003  
LOCUS Synthetic construct Homo sapiens CDC-like kinase 4 (CLK4) mRNA,  
DEFINITION partial cds.  
ACCESSION AY335726  
VERSION AY335726.1 GI:33304090  
KEYWORDS FLJ CDNA.  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1 (bases 1 to 1446)  
AUTHORS Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D.,  
Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepsen,D., Harlow,E.,  
Labaer,J. and Brizuela,L.  
Cloning of human full-length CDS FLEXGene kinases in  
recombinational vector system  
TITLE Unpublished  
JOURNAL 2 (bases 1 to 1446)  
REFERENCE Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D.,  
AUTHORS Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepsen,D., Harlow,E.,  
Labaer,J. and Brizuela,L.  
TITLE Direct Submission  
JOURNAL Submitted (02-JUL-2003) Biological Chemistry and Molecular  
Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,  
Cambridge, MA 02141-2023, USA  
COMMENT This CDS clone is a part of a collection of human full-length  
expression clones generated by Harvard Institute of Proteomics.



Altechul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerbach, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1914)

Strausberg, R.

Direct Submission

Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

Contact: [amadan@systemsbiology.org](mailto:amadan@systemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 18 Row: c Column: 18

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6671765.

#### FEATURES

Location/Qualifiers

1..1914

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="MGC:13871 IMAGE:3982920"

/tissue\_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMV."

/clone\_lib="NCI CGAP Lu29"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..1914

/gene="Clk4"

/codon\_start=1

/product="CDC like kinase 4"

/protein\_id="AAH12675.1"

/db\_xref="GI:15215131"

/db\_xref="LocusID:12750"

/db\_xref="MGI:1098551"

/translation="MRHSKRTCPDWSRSGHSGSHKRSRSHSSQTQNRHCK PHQFQSDCHYLEARCLNDRYDRYIDRNDYCEGYVPRHYRHDVESTYRHCSSVRSRSHSSPKRNRCPASHQSHRSRSHRSIEDDEGLHICQSGDLRARYEIVDTLGEAGFGKVECDHGMGLHVAVKIVQNVGRREARSEIQVLEHLNSTDPNS

#### gene

#### CDS

VFRVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPFOIDHIRMAYQICQISINF LHHKLTHTDLKPNILPVKSDYVYKNSKMRDERTKNTDQIVDFGSAITDEHHI LSTVSRHYRAPEVLALGWSQPCVMSIGILIEYLGFTVFOVDFSKHLLAMMERI LGPIPAHMIQKTRRKVYFHHNQLDWDHSSAGRYVRRRCCKPLKBFMLCHDEHEKLFDF LVRMLEYDPPARRITLDEALQHPFDLLKRR"

#### ORIGIN

Query Match 56.6%; Score 1277.6; DB 10; Length 1914;  
Best Local Similarity 87.5%; Pred. No. 6.5e-227;  
Matches 1418; Conservative 0; Mismatches 200; Indels 3; Gaps 2;

QY 14 GTCATTATTAGAACGAGTCTTGTGATGACGAGATATTCCGACCGGAGATAGCTTG 73  
DB 261 GTCACATATTAGAACGAGATGCTTGAATGAGAGAGATTATCGGACCGGAGATACAT 320  
QY 74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTCTAGACATATTATCAGAGACAT 133  
DB 321 ATGAATACAGGAATGACTACTGTGAGAGGATATGTTCTCAAGACATATTACCATGAGACGTTG 380  
QY 134 AAAGCGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAGAGACAGTCTCTA 193  
DB 381 AAAGCACCTTACCGGATCCATTCAGTAAATCTTCAGTCCGACGAGGAGAGACAGCCTTA 440  
QY 194 AAAGGAGCGCAATAGACATCTGTTCAAGTCACTAGTTCAGTTCGAGAGCCACCGAAGA 253  
DB 441 AGAAGAACGTAATAGACCTGTGCAAGTCACTAGTTCGAGAGCCACCGAAGA 500  
QY 254 AAAGATCCAGGAGTATGAGGATGATGAGGAGGTCACCTGATCTGTCAGAGTGAGACG 313  
DB 501 AAAGATCCAGGAGTATGAGGATGATGAGGAGGTCACCTGATCTGTCAGAGTGAGACG 560  
QY 314 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGCCCTTTGGCAAAGTTG 373  
DB 561 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTAGGTGAAGAGCCCTTTGGCAAAGTTG 620  
QY 374 TAGAGTGCAATTGATCATGGCATGGCATGGCATGGCATGGATGAGAGTGAAGTTCGTAAGAA 433  
DB 621 TAGAGTGCAATTGATCATGGCATGGCATGGCATGGCATGGATGAGAGTGAAGTTCGTAAGAA 680  
QY 434 TAGCGCGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCACTTTAAATAGTA 493  
DB 691 TAGCGCGTTACCGGAGGAGCAGCTCGTTCTGAATCCAAAGTATTGGAGCACTTTGAACAGA 740  
QY 494 CTGATCCCAATAGTGTCTTCGAGTGTGTCCAGATGCTAGAGTGGTTGATCATCATGTCGTC 553  
DB 741 CTGACCCCAACAGTGTCTTCGAGTGTGTCCAGATGCTAGAGTGGTTGATCATCATGTCGTC 800  
QY 554 ATGTTTGTATTGTGTTTGAACACTCTGGGACCTTAGTACTTACATTTTCAATTAAGAAAAA 613  
DB 801 ATGTTTGTATTGTGTTTGAAGCTGTGGGACCTTAGTACTTACATTTTCAATTAAGAAAAA 860  
QY 614 GCTTTCTGCCAATTTCAAATTCACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 673  
DB 861 GTTTTCTGCCAATTTCAAATTCACATCAGGAGAAATGGCTTATCAGATCTGCCAGTCTA 920  
QY 674 TAAATTTTTTATCATATAATTAATTAACCCATACAGATCTCGAAGCCTGAAATATTTTGT 733  
DB 921 TAAATTTTTTATCATATAATTAATTAACACACAGGACCTTAAACCTGAAATATTTTAT 980  
QY 734 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTTAAATGAAGCTGATGAACCCACAC 793  
DB 981 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTTAAATGAAGCTGATGAACCCACAC 1040  
QY 794 TGAAGAACACAGATATCAAGTGTGTTGAGTGTGGAAGTGCAACCTATCATGATGATGAACATC 853  
DB 1041 TGAAGAACACAGATATCAAGTGTGTTGAGTGTGGAAGTGCAACCTATCATGATGATGAACATC 1100  
QY 854 ACAGTACTTTTGTGCTCTACCCCGGACATACAGAGTCCCGAGGCTCAATTTTGGCTTTAGGTT 913  
DB 1101 ATAGTACTTTTGTGCTCCACAGGACACTACAGGCTCCAGAGGCTCAATTTTGGCTCTAGGTT 1160  
QY 914 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGATCTTCTTATTTGAATATTACCTTGGTT 973

```

Db 1161 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGCTGCAATCTTATTGAGTACTACCTTGGGT 1220
QY 974 TCACAGTCTTTTCAGACTCATGATAGTAAGACCACTCGCAATGATGAAAGTATTAG 1033
Db 1221 TCACAGTCTTTTCAGACCCACCATAGTAAAGAGCACTCGCAATGATGAGCGGATCTTAG 1280
QY 1034 GACCCATACCAACAACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTCACCATACCC 1093
Db 1281 GACCCATCCAGCACATATGATCCAGAGACAAGAAAGACGATATTTCACCATACCC 1340
QY 1094 AGCTAGATGGGATGAAACACAGTCTTCTGCTGATGATGTTAGGAGACGCTGCAACCGT 1153
Db 1341 AGCTAGATGGGACGAGCATAGTTGAGCTGGGAGATGTTAGGAGACGCTGCAACCGT 1400
QY 1154 TGAAGGATTTATGCTTTGTCATGATGAAGACATGAGAAATCTTTGACCTGTTCCAA 1213
Db 1401 TAAAGGAATTTATGCTGTCATGACGAAGAGCATGAGAAAGTCTTTGACCTGTTCCAA 1460
QY 1214 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCTT 1273
Db 1461 GAATGTTGAGATGATGACCCAGGAGAGATCACCTTGGATGAAGCATTTGCAGCACCTT 1520
QY 1274 TCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTATATATCTTCT 1333
Db 1521 TCTTTGACTTATTAAGAAAGAAATG-AGTGGAGTCAAGGCTCTTCTGCTACTTCTCTA 1579
QY 1334 AGAGAGATTTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1393
Db 1580 GGAGCAGTTACTTCAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1639
QY 1394 AAATTTATTTGACAGTTAAAGTGAATATTTGATGTTTGTATCAATAGCATATTAAC 1453
Db 1640 AAATTTATTTGACAGTTAAATGTAATATCTGATGTTTGTATCAATAGCATATTAAC 1699
QY 1454 TTGTTAAGCAAGTATGCTTGAATGAATGATAGAAATTAATATTTTCTTTT 1512
Db 1700 TTGTTAAGTATGCTTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1759
QY 1513 -TGAATTTACCATTTTAAATACCTTTGAAATATCTTTGTCAGTGAATATGAT 1571
Db 1760 ATGTTAATGCGCTTTTAAAGCTTTTGAAGCTTTTGAAGCTTTTGAAGCTTTTGAAGCT 1819
QY 1572 TGATCTTGCCTTTTGTACATGAGGTGACCTCTGAAGTGAATTTTGTGATTAAGGAA 1631
Db 1820 TGGTCCATCTTTGTACATGAGGTGACCTCTGAAGTGAATTTTGTGATTAAGGAA 1879
QY 1632 A 1632
Db 1880 A 1880

RESULT 13
AX961894 1881 bp DNA linear PAT 14-JAN-2004
LOCUS Sequence 105 from Patent WO03104277.
DEFINITION AX961894
ACCESSION AX961894
VERSION AX961894.1 GI:40881304
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
AUTHORS Stat6 activation gene
TITLE Patent: WO 03104277-A 105 18-DEC-2003;
JOURNAL Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
Location/Qualifiers
1..1881
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
90..1535
CDS

```

```

/note="unnamed protein product"
/codon_start=1
/protein_id="CAP06456.1"
/db_xref="GI:40881305"
/translation="MRHSKTHCPDWDSDRESWGHSYSGSHKRRKSRSHSSTQENRHK
PHQFDSCHYLEARCLNBRDYDRYIDYENDYCEGVYPRHYHRSYRHS
KSSVRSRSPKRRNRPCASHQSHSRKRSRSEDDEEHLICQSGVLRARVE
IVDTLGGAGTKVVCIDHGMGLHVAIKVKNVRYREARSEIQVLEHNSQSPNS
VFCVOMLEWFDHGHVCIVFELLGLSTYDFIKENSLFQIDHIRMVAVQICQINF
LHNKLPHDLKPNILLFKSDYVVKVNSKMRDERTLKNTDIKVDFGSAFYDEHH
STLVSTRHRAPEVILALHNSQPCDWSISGILLEYLIGTVFOTHSKHLAMERI
LGP1PAHMIQKTRKRYPHNLDWDEHSSAGRYVRRRCKPKEFLCHDEHEKFLFD
LVRRMLEYDPARRITLDELQHPFDLLKRX"
ORIGIN
Query Match 56.6%; Score 1277.4; DB 6; Length 1881;
Best Local Similarity 87.9%; Pred. No. 7.1e-227; Indels 3; Gaps 2;
Matches 1413; Conservative 0; Mismatches 192;
QY 14 GTCATTATTAGAACCAAGTCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG 73
Db 250 GTCATTATTAGAACCAAGTCTTGAATGAGAGAGATTATCGGGACCGGAGATACATTG 309
QY 74 ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTATCACAGACATTG 133
Db 310 ATGAATACAGAAATGACTACTGTGCAAGGATATGTTTCCAAGACATTATCCATAGACGTTG 369
QY 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGCGAGGAGAGAGAGTCCCTA 193
Db 370 AAAGCATTTACCGGATCCATTGCGAGTAAATCCTCAGTCAGGACGAGGAGAGAGAGTCCCTA 429
QY 194 AAAGAAAGCGCAATAGACACTGTTTCAAGTCAATCAGTCACGTTTCAAGAGCACCCGAAAGGA 253
Db 430 AGAAAGCGTAAATAGACCTGTGCAAGTCAATCAGTCACGTTTCAAGAGCACCCGAAAGGA 489
QY 254 AAAGATCCAGAGTATAGAGATGATGAGAGGGTCACTGATCTGTCAAGAGTGAGAGCG 313
Db 490 AAAGATCCAGAGTATAGAGATGATGAGAGGGTCACTGATCTGTCAAGAGTGAGAGCG 549
QY 314 TTCTAAGAGCAAGATATGAATCTGGACACTTTTGGGTGAAGGAGCGCTTTGCGCAAGTTG 373
Db 550 TTCTAAGAGCAAGATATGAATCTGGACACTTTTGGGTGAAGGAGCGCTTTGCGCAAGTTG 609
QY 374 TAGAGTCAATTGATCATGGCATGGATGCGCATGATGATGATGATGATGATGATGATGATGATGAT 433
Db 610 TAGAGTCAATTGATCATGGCATGGATGCGCTTACATGATGATGATGATGATGATGATGATGATGAT 669
QY 434 TAGGCGGTTACCGTGAAGCAGCTGTTCAAGAAATCCAAAGTATTAGAGCACTTAATAGTA 493
Db 670 TAGGCGGTTACCGGAGGAGCAGCTGTTCTGAAATCCAAAGTATTAGAGCACTTGAACAGCA 729
QY 494 CTGATCCCAATAGTGTCTTCCGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGT 553
Db 730 CTGATCCCAATAGTGTCTTCCGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGT 789
QY 554 ATGTTTGTATTGTTGTTGAACTACTGGAAGTCTTAGTACTTACGATTTCAATTAAGAAAAACA 613
Db 790 ATGTTTGTATTGTTGTTGAGCTGCTGGGACTTAGTACTTACCTATGATTTTATTAAGAAAAATA 849
QY 614 GCTTTCTGCCATTTCAAAATGACCAATCAGCAGATGCGGATGATGATGATGATGATGATGATGATGAT 673
Db 850 GCTTTCTGCCATTTCAAAATGATCAGATCAGGCAATGGCTTATCAGATCTGCCAGTCTA 909
QY 674 TAAATTTTTCATCATATAATAATTAACCCATACAGATCTCAAGGCTCAAGGCTCAAGGCTCAAGGCT 733
Db 910 TAAATTTTTCATCATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 969
QY 734 TTGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 793
Db 970 TTGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
QY 794 TGAAGAAACAGATATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 853

```





|                            |      |  |      |
|----------------------------|------|--|------|
| QY                         | 1166 | TGCTTTGTGTCATGATGAAGAAACATGAGAAACTGTTTGACCTGGTTCGAAGAATGTTAGAAAT | 1225 |
| Db                         | 729  | TGCTTTGTGTCATGATGAAGAAACATGAGAAACTGTTTGACCTGGTTCGAAGAATGTTAGAAAT | 788  |
| QY                         | 1226 | ATGATCCAACTCAAAGAAATACCTTTGGATGAAGCAATGCAGATCCTTTCTTGACTTAT      | 1285 |
| Db                         | 789  | ATGATCCAACTCAAAGAAATACCTTTGGATGAAGCAATGCAGATCCTTTCTTGACTTAT      | 848  |
| QY                         | 1286 | TAAAAAGAAATGAATGGGAATCAGTGGTCTTACTATATATCTCTAGAGAGATTAC          | 1345 |
| Db                         | 849  | TAAAAAGAAATGAATGGGAATCAGTGGTCTTACTATATATCTCTAGAGAGATTAC          | 908  |
| QY                         | 1346 | TTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGT       | 1405 |
| Db                         | 909  | TTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGT       | 968  |
| QY                         | 1406 | ACAGTTAAGTGAATATATGATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAG        | 1465 |
| Db                         | 969  | ACAGTTAAGTGAATATATGATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAG        | 1028 |
| QY                         | 1466 | TATGCTCTTGATATGATGCAATAGAAAATTAATTTTCTTTTGTAAATTAACCAAT          | 1525 |
| Db                         | 1029 | TATGCTCTTGATATGATGCAATAGAAAATTAATTTTCTTTTGTAAATTAACCAAT          | 1088 |
| QY                         | 1526 | TTTAAATACCTTTGAAATAT - CCTTTGTGTCAGTGAATAATGATGATCTTGCCCTTT      | 1584 |
| Db                         | 1089 | TTTAAATACCTTTGAAATATACCTTTGTGTCAGTGAATAATGATGATCTTGCCCTTT        | 1148 |
| QY                         | 1585 | TGTACATGGAGTCACTCTGAAGTGATTTTTTTTGAGTAAAGAAATCTTGACTA            | 1641 |
| Db                         | 1149 | TGTACATGGAGTCACTCTGAAGTGATTTTTTTTGAGTAAAGAAATCTTGACTA            | 1205 |
| RESULT 15                  |      |  |      |
| AR380154                   |      |  |      |
| LOCUS                      |      |  |      |
| DEFINITION                 |      |  |      |
| SEQUENCE                   |      |  |      |
| ACCESSION                  |      |  |      |
| VERSION                    |      |  |      |
| KEYWORDS                   |      |  |      |
| SOURCE                     |      |  |      |
| ORGANISM                   |      |  |      |
| REFERENCE                  |      |  |      |
| AUTHORS                    |      |  |      |
| TITLE                      |      |  |      |
| JOURNAL                    |      |  |      |
| FEATURES                   |      |  |      |
| ORIGIN                     |      |  |      |
| Query Match                |      |  |      |
| Best Local Similarity      |      |  |      |
| Matches 1191; Conservative |      |  |      |
| QY                         | 932  | TTTGAGCATAGGTTGCAATCTTATTAATATTAACCTTTGGTTTTCAGTCTTTGAGCTC       | 991  |
| Db                         | 262  | TGTAGGCATAGGTTGCAATCTTATTAATATTAACCTTTGGTTTTCAGTCTTTGAGCTC       | 321  |
| QY                         | 992  | ATGATAGTAAAGAGCACCTGGCAATGATGGACGAATATTAGGACCCATACCACAAACACA     | 1051 |
| Db                         | 322  | ATGATAGTAAAGAGCACCTGGCAATGATGGACGAATATTAGGACCCATACCACAAACACA     | 381  |
| QY                         | 1052 | TGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACCCAGCTAGATTGGGATGAAC  | 1111 |
| Db                         | 382  | TGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACCCAGCTAGATTGGGATGAAC  | 441  |
| QY                         | 1112 | ACAGTTCTGCTGGTATATGTTAGGACGCTGCAACCCGTTGAAGGAATTTATGCTTT         | 1171 |

Search completed: March 16, 2005, 14:31:59  
Job time : 9712.74 secs

|    |      |   |      |
|----|------|---|------|
| Db | 442  | ACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTATGCTTT | 501  |
| QY | 1172 | GTCATGATGAAGAAACATGAGAAACTGTTTGACCTGGTTCGAAGAATGTTAGAAATGATC  | 1231 |
| Db | 502  | GTCATGATGAAGAAACATGAGAAACTGTTTGACCTGGTTCGAAGAATGTTAGAAATGATC  | 561  |
| QY | 1232 | CAACTCAAGAAATTAACCTTTGGATGAAGCAATGCAGATCCTTTCTTGACATTTAAAAA   | 1291 |
| Db | 562  | CAACTCAAGAAATTAACCTTTGGATGAAGCAATGCAGATCCTTTCTTGACATTTAAAAA   | 621  |
| QY | 1292 | AGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG  | 1351 |
| Db | 622  | AGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG  | 681  |
| QY | 1352 | CTGTTGTCAGTCACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGTACAGTT    | 1411 |
| Db | 682  | CTGTTGTCAGTCACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGTACAGTT    | 741  |
| QY | 1412 | AAGTGAATAATTTGATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAGTATGTT    | 1471 |
| Db | 742  | AAGTGAATAATTTGATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAGTATGTT    | 801  |
| QY | 1472 | CTTGATTAATGCAATAGAAAATTAATTTTCTTTTGTAAATTAATTTTCTTTTGTAA      | 1530 |
| Db | 802  | CTTGATTAATGCAATAGAAAATTAATTTTCTTTTGTAAATTAATTTTCTTTTGTAA      | 861  |
| QY | 1531 | ATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGATGATCTTGCCCTTTTGTACA     | 1590 |
| Db | 862  | ATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGATGATCTTGCCCTTTTGTACA     | 921  |
| QY | 1591 | TGGAGTCACTCTGAAAGTGAATTTTTTTTGAGTAAAGAAATCTTGACTATTTATATT     | 1650 |
| Db | 922  | TGGAGTCACTCTGAAAGTGAATTTTTTTTGAGTAAAGAAATCTTGACTATTTATATT     | 981  |
| QY | 1651 | CTTAAAGCAATTTCTTTTATATCTCAAAATTTAGAACTTTAAAGTCTTTTCTTC        | 1710 |
| Db | 982  | CTTAAAGCAATTTCTTTTATATCTCAAAATTTAGAACTTTAAAGTCTTTTCTTC        | 1041 |
| QY | 1711 | TGTAATTTGTCGAAGGGTGATTTATTAATTAACCTAGATAAGCAGTACTAGAAACAAA    | 1770 |
| Db | 1042 | TGTAATTTGTCGAAGGGTGATTTATTAATTAACCTAGATAAGCAGTACTAGAAACAAA    | 1101 |
| QY | 1771 | CTCAGAAATGTTTCTGTTAGAAATTTCTAATTAATTTTAAAGTGTGTTATCTTTTTCAT   | 1830 |
| Db | 1102 | CTCAGAAATGTTTCTGTTAGAAATTTCTAATTAATTTTAAAGTGTGTTATCTTTTTCAT   | 1161 |
| QY | 1831 | GGGTGATGTCAGGGTGATAACCAAGCAATTCATGGAAGGCGATGCAATTTGTCATTGTA   | 1890 |
| Db | 1162 | GGGTGATGTCAGGGTGATAACCAAGCAATTCATGGAAGGCGATGCAATTTGTCATTGTA   | 1221 |
| QY | 1891 | CAGTTTGTTTTAAATAAACCACATACACATTTTATTTAAGATTTAAATCTAACTGGAAGT  | 1950 |
| Db | 1222 | CAGTTTGTTTTAAATAAACCACATACACATTTTATTTAAGATTTAAATCTAACTGGAAGT  | 1281 |
| QY | 1951 | CAGTTTGGAAAAATGGAATTTCCAAAGTATGTTGGTGAAGTCAAGATTAATAAATAGAAA  | 2010 |
| Db | 1282 | CAGTTTGGAAAAATGGAATTTCCAAAGTATGTTGGTGAAGTCAAGATTAATAAATAGAAA  | 1341 |
| QY | 2011 | TTCTGATGAGAGGTTTTCAGTTTAAATCAAGTCTTTAGGAGTCTTAACATTTGCCAGC    | 2070 |
| Db | 1342 | TTCTGATGAGAGGTTTTCAGTTTAAATCAAGTCTTTAGGAGTCTTAACATTTGCCAGC    | 1401 |
| QY | 2071 | ATCTGTTTATCAAAATGACATAAATACGTAACCTATAAAGATTAAGTTTATTAAT       | 2125 |
| Db | 1402 | ATCTGTTTATCAAAATGACATAAATACGTAACCTATAAAGATTAAGTTTATTAAT       | 1456 |



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 23:59:12 ; Search time 41 Seconds  
(without alignments)  
1044.304 Million cell updates/sec

Title: US-10-801-671-2

Perfect score: 2410

Sequence: 1 MCIPLEASHSVEEDTHPSHY.....QRITLDEALQHFFDLKKK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_79.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 1887  | 78.3        | 484    | 2 S53641 | protein kinase clk |
| 2          | 1837  | 76.2        | 483    | 2 A39676 | protein kinase STY |
| 3          | 1727  | 71.7        | 454    | 2 A38643 | protein kinase EC  |
| 4          | 1417  | 58.8        | 499    | 2 S3637  | protein kinase clk |
| 5          | 1321  | 54.8        | 490    | 2 S3639  | protein kinase clk |
| 6          | 1317  | 54.6        | 490    | 2 S70352 | protein kinase Dar |
| 7          | 1222  | 50.7        | 517    | 2 A54099 | protein kinase Dar |
| 8          | 965.5 | 40.1        | 903    | 2 T19209 | probable protein k |
| 9          | 876   | 36.3        | 306    | 2 I49068 | protein kinase STY |
| 10         | 848.5 | 35.2        | 427    | 2 T05560 | protein kinase AFC |
| 11         | 839   | 34.8        | 431    | 2 T04125 | protein kinase PKI |
| 12         | 830   | 34.4        | 467    | 2 S71169 | protein kinase, 54 |
| 13         | 790   | 32.8        | 400    | 2 T04460 | protein kinase AME |
| 14         | 767   | 31.8        | 575    | 2 JC7794 | lammer kinase homo |
| 15         | 767   | 31.8        | 690    | 2 T38052 | probable protein k |
| 16         | 718.5 | 25.8        | 737    | 1 S64767 | probable serine/th |
| 17         | 530   | 22.0        | 543    | 2 S62456 | probable serine-th |
| 18         | 524   | 21.7        | 1087   | 2 S58147 | protein kinase - f |
| 19         | 523   | 21.7        | 508    | 2 T22440 | hypothetical prote |
| 20         | 523   | 21.7        | 817    | 2 T22442 | hypothetical prote |
| 21         | 504.5 | 20.9        | 589    | 2 JG0195 | protein kinase DYT |
| 22         | 478.5 | 19.9        | 629    | 2 JG0195 | protein kinase DYT |
| 23         | 477   | 19.8        | 1157   | 2 C96761 | hypothetical prote |
| 24         | 474   | 19.7        | 948    | 2 T24445 | hypothetical prote |
| 25         | 472.5 | 19.6        | 754    | 2 JC4898 | Down-syndrom-crit  |
| 26         | 448.5 | 18.6        | 1189   | 2 T17088 | homeodomain-intera |
| 27         | 443.5 | 18.4        | 1155   | 2 B96761 | probable protein k |
| 28         | 441   | 18.3        | 1192   | 2 T17089 | homeodomain-intera |
| 29         | 437   | 18.1        | 642    | 2 T45904 | protein kinase-lik |

probable protein k  
serine/threonine p  
protein kinase YAK  
hypothetical prote  
protein kinase YAK  
homeodomain-intera  
mRNA splicing-asso  
protein B0464.5a l  
hypothetical prote  
hypothetical prote  
probable protein k  
protein kinase hom  
protein B0464.5c l  
protein kinase (EC  
serine protein kin  
mitogen-activated

30 435 18.0 570 2 E84825  
31 432 17.9 1191 2 T14154  
32 431.5 17.9 807 2 A32582  
33 430 17.8 821 2 T34232  
34 429 17.8 1457 2 T14577  
35 419 17.4 1209 2 T14357  
36 406.5 16.9 477 2 T11720  
37 401 16.6 1093 2 F88556  
38 398 16.5 775 2 T21259  
39 392 16.3 1087 2 S28282  
40 389 16.1 440 2 C84553  
41 380.5 15.8 438 2 T04655  
42 367.5 15.2 782 2 E88556  
43 343 14.2 353 2 S28548  
44 338.5 14.0 523 2 T47560  
45 334 13.9 361 2 T51943

ALIGNMENTS

RESULT 1  
S53641  
protein kinase clk1 (EC 2.7.1.1) - human  
C:Species: Homo sapiens (man)  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C:Accession: S53641  
R:Hanes, J.; von der Kammer, H.; Klaudiny, J.; Scheit, K.H.  
J. Mol. Biol. 244, 665-672, 1994  
A:Title: Characterization by cDNA cloning of two new human protein kinases. Evidence by  
A:Reference number: S53637; MUID:95082033; PMID:7990150  
A:Accession: S53641  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-484 <HAN>  
A:Cross-references: UNIPROT:P49759; GB:L29219; NID:G632963; PIDN:AAAG1480.1; PID:G632964  
C:Superfamily: human protein kinase clk1; protein kinase homology  
C:Keywords: alternative splicing; phosphotransferase  
F:159-430/Domain: protein kinase homology <KIN>

Query Match 78.3%; Score 1887; DB 2; Length 484;

Best Local Similarity 82.1%; Pred. No. 2.1e-83;

Matches 352; Conservative 30; Mismatches 45; Indels 2; Gaps 2;

QY 18 SHYLEARSLNERDYRRYDEYNDYCEGVVPRHYHRDIESGYRIHCKSSVRRSSP 77

DB 54 SHYLESRLNEKYHSRYIDEYNDYQGCPEGRQDRHSRYQNHSSKSSGRSSY 113

QY 78 KRK-RNRHCSSH-QSRKSHRRKRSRSTEDDEEGLICQSGDVLARARYEIVDTLGEAGF 135

DB 114 KSKRIHHSTSHRRSHGSKSHRRKTRSVEDDEEGLICQSGDVLARARYEIVDTLGEAGF 173

QY 136 KVCEDHGMGMVAVKIVKNGVRYEAAARSEIQVLEHLNSTDPNSVFCVQMLEWDH 195

DB 174 KVCEDHKGGRVAVKIVKNDRYCEAAARSEIQVLEHLNTDPNSTFRCVQMLEWFEH 233

QY 196 HGHYCVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQINFLHNKLTHTDLKPE 255

DB 234 HGHYCVFELLGLSTYDFIKENGLPFLDHRKMAVQICKSVNLFHNSKLTHTDLKPE 293

QY 256 ILFKVSDYVVKYNSKMKRDERTLNKTDIKVDFGATYDDDEHSTLVSTRYRADEVILA 315

DB 294 ILFKVSDYVVKYNSKMKRDERTLNKTDIKVDFGATYDDDEHSTLVSTRYRADEVILA 353

QY 316 LGWSQPCDVWSIGCILLEYLGFVTFQTHDSKEHLAMMERILGPTPOHMIQKTRKRYFH 375

DB 354 LGWSQPCDVWSIGCILLEYLGFVTFQTHDSKEHLAMMERILGPTPOHMIQKTRKRYFH 413

QY 376 HNOQDWEHSSAGRYVRRRKPLKEFMLEKFDLVRMLEYDPTQRTITDEALQ 435

DB 414 HNRLDWEHSSAGRYVRRRKPLKEFMLEKFDLVRMLEYDPTQRTITDEALQ 473

QY 436 HPFFDLLKK 444

C;Date: 21-Feb-1992 #sequence\_revision 21-Feb-1992 #text\_change 09-Jul-2004

C;Accession: A38643  
R;Johnson, K.W.; Smith, K.A.  
J. Biol. Chem. 266, 3402-3407, 1991  
A;Title: Molecular cloning of a novel human cdc2/CDC28-like protein kinase.  
A;Reference number: A38643; MUID:91139618; PMID:1704889  
A;Accession: A38643  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-454 <JOH>  
A;Cross-references: UNIPROT:Q9NRL6; GB:M59287  
C;Genetics:  
A;Gene: GDB:CLK  
A;Cross-references: GDB:126862  
A;Map position: 7q31-7q31  
C;Superfamily: human protein kinase clk1; protein kinase homology  
C;Keywords: phosphotransferase  
F;129-400/Domain: protein kinase homology <KIN>

Query Match 71.7%; Score 1727; DB 2; Length 454;  
Best Local Similarity 76.6%; Pred. No. 8.5e-76;  
Matches 327; Conservative 26; Mismatches 46; Indels 28; Gaps 3;  
Qy 18 SHYLEARSLNERDYRRYVDEYRNDYCEGVPRHYHRDIESGYRIHCSKSSVRSRSP 77  
Db 54 SHYLESRINEKYHSRRYIDEYRNDYTCQCEPGHRRDHSRYQNHSSKSSGRSSY 113  
Qy 78 KRKNRHCSSHQSRSKSHRRKRSRSIEDDEBEGHLICQSGDVLRYRVEIVDTLGEAGFKV 137  
Db 114 KSKRIH-----HSTSHRSH-----GD-----EIVDTLGEAGFKV 145  
Qy 138 VECIDHGMGMHVAKIVKNGVRYREARAEIQLVLEHLNSTDPNSVRCQMLEWFDHGG 197  
Db 146 VECIDHKGAGRHVAKIVKNDYCEAARSEIQVLEHLNTTDPNSTPRCQMLEWFFHHG 205  
Qy 198 HVCIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSFNHLHNKLTHTDLKPENIL 257  
Db 206 HICIVFELLGLSTYDFIKENGLFPLDHIKMAQYICKSVNHLNKLHTDLKPENIL 265  
Qy 258 FVKSIDYVVKNSKMKRDETLKNTDIKVDPGSAFYDDEHSTLVSTRHYRAPEVILALG 317  
Db 266 FVQSDYTEAYNPKIKRDETLINPDIKVDPGSAFYDDEHSTLVSTRHYRAPEVILALG 325  
Qy 318 WSPQCDVWSICILIEYVLTGTFQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHN 377  
Db 326 WSPQCDVWSICILIEYVLTGTFVPTTHDSKEHLAMMERILGPKHMLQKTRKRYFHD 385  
Qy 378 OLDWDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRLMEYDPTQRTITLDEALQHP 437  
Db 386 RLDWDEHSSAGRYVSRACKPLKEFMLSQDVEHERLFDLIQKMLEYDPAKRITLREALKHP 445  
Qy 438 PFDLLKK 444  
Db 446 PFDLLKK 452

RESULT 4  
SS3637  
protein kinase clk2, long splice form (EC 2.7.1.1-) - human  
C;Species: Homo sapiens (man)  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C;Accession: S53637; T08825  
R;Hanes, J.; von der Kammer, H.; Klaudiny, J.; Scheit, K.H.  
J. Mol. Biol. 244, 665-672, 1994  
A;Title: Characterization by cDNA cloning of two new human protein kinases. Evidence by  
A;Reference number: S53637; MUID:95082033; PMID:7990150  
A;Accession: S53637  
A;Molecule type: mRNA  
A;Residues: 1-499 <HAN>  
A;Cross-references: UNIPROT:P49760; GB:I29218; NID:G632967; PIDN:AAA61482.1; PID:G632968  
R;Winfield, S.L.; Tayebi, N.; Martin, B.M.; Gims, E.I.; Sidransky, E.  
Genome Res. 7, 1020-1026, 1997  
C;Species: Homo sapiens (man)  
A;Title: Identification of three additional genes contiguous to the glucocerebrosidase 1c

C;Date: 08-Nov-1991 #sequence\_revision 08-Nov-1991 #text\_change 09-Jul-2004

C;Accession: A39676; S13364  
R;Howell, B.W.; Afar, D.E.H.; Lew, J.; Douville, E.M.J.; Icelly, P.L.E.; Gray, D.A.; Bell  
Mol. Cell. Biol. 11, 568-572, 1991  
A;Title: STY, a tyrosine-phosphorylating enzyme with sequence homology to serine/threonine  
A;Reference number: A39676; MUID:91094875; PMID:1986248  
A;Accession: A39676  
A;Molecule type: mRNA  
A;Residues: 1-483 <HOW>  
A;Cross-references: UNIPROT:P22518; GB:M38381; NID:G201070; PIDN:AAA40151.1; PID:G201071  
R;Ben-David, Y.; Letwin, K.; Tannock, L.; Bernstein, A.; Pawson, T.  
EMBO J. 10, 317-325, 1991  
A;Title: A mammalian protein kinase with potential for serine/threonine and tyrosine pho  
A;Reference number: S13364; MUID:91122038; PMID:1825055  
A;Accession: S13364  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-378, P', 380-483 <BEN>  
A;Note: the sequence from Fig. 2 is inconsistent with that shown in Fig. 1 in having 448  
C;Superfamily: human protein kinase clk1; protein kinase homology  
C;Keywords: phosphotransferase; serine/threonine-specific protein kinase  
F;158-429/Domain: protein kinase homology <KIN>

Query Match 76.2%; Score 1837; DB 2; Length 483;  
Best Local Similarity 79.7%; Pred. No. 5.1e-81;  
Matches 349; Conservative 32; Mismatches 51; Indels 6; Gaps 4;  
Qy 9 HSVEEDTHPSHYLEARSLNERDYRRYVDEYRNDYCEGVPRHYHRDIESGYRIHCSKS 68  
Db 48 HSKTTD--SYLESRSINEKAYHSRRYVDEYRNDY-GYEPGHPYGPGRSGYQMHSSKS 103  
Qy 69 SVRSRRSPKPK-RN-HCSHQSRKSHRRKRSRSIEDDEBEGHLICQSGDVLRYRVEIV 126  
Db 104 SGRSGRSSYKSHRSRHTSQHSHGKSHRRKRSRSVEDDEBEGHLICQSGDVLRYRVEIV 163  
Qy 127 DTLGEAGFKVVECIDHGMGMHVAKIVKNGVRYREARAEIQLVLEHLNSTDPNSVRC 186  
Db 164 DTLGEAGFKVVECIDHKGVRRAVAVKIVNDYCEAARSEIQVLEHLNTTDPHSTPRC 223  
Qy 187 VQMLEWFDHGHVCIIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSFNHLHNKL 246  
Db 224 VQMLEWFEHRGHICIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICKSVNHLNKL 283  
Qy 247 THTDLPENILFVKSIDYVVKNSKMKRDETLKNTDIKVDPGSAFYDDEHSTLVSTRH 306  
Db 284 THTDLPENILFVKSIDYTEAYNPKIKRDETLINPDIKVDPGSAFYDDEHSTLVSTRH 343  
Qy 307 YRAPEVILALGWSQCDVWSIGCILIEYVLTGTFQTHDSKEHLAMMERILGPIQHMIOQ 366  
Db 344 YRAPEVILALGWSQCDVWSIGCILIEYVLTGTFVSTHDSREHLAMMERILGPKHMIQ 403  
Qy 367 KTRKRYFHNLQDWDDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRLMEYDPTQ 426  
Db 404 KTRKRYFHNLQDWDDEHSSAGRYVRRCKPLKEFMLSQDAEHELLFDLIGKMLEYDPAK 463  
Qy 427 RITLDEALQHPFDLLKK 444  
Db 464 RITLKEALKHPFFYPLKK 481

RESULT 3  
A38643  
protein kinase (EC 2.7.1.37) cdc2/cdc28-like - human  
C;Species: Homo sapiens (man)

A:Reference number: Z16482; MUID:97474796; PMID:9331372  
A:Accession: T08825  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-499 <WIN>  
A:Cross-references: GB:AF023268; NID:G2564910; PIDN:AA654911; PID:G2564911  
C:Comment: The short splice form of this protein (see PIR:S53638) lacks the protein kinase domain  
C:Genetics:  
A:Gene: clk2  
C:Function:  
A:Description: Phosphorylates Ser, Thr, and Tyr residues on proteins in the spliceosomal  
C:Superfamily: human protein kinase clk1; protein kinase homology  
C:Keywords: alternative splicing; phosphotransferase  
F:161-440/Domain: protein kinase homology <KIN>

Query Match 58.8%; Score 1417; DB 2; Length 499;  
Best Local Similarity 59.7%; Pred. No. 5, 5e-61;  
Matches 259; Conservative 69; Mismatches 98; Indels 8; Gaps 3;

QY 18 SHYLEARS-----LNERDYRRRYVDEY-RNDYCEGYVPRHYRDIESGYRHCKSVSR 71  
Db 50 SYHVSRSYSDRRSDRRYDRYCGSYRNDYSDRDGDYVDTYRHSYEQRENSYR 109  
QY 72 SRRSPK--RKRNHCSHQSRKSHRRKRSIESIEDDEGHLICQSGDVLARVEIVDTL 129  
Db 110 SQRSRRKRRRRRSRTFSRSSQHSRRRAKSVEDDAEGHLIYHVGDLQERYEIVSTL 169  
QY 130 GEGAFKGVVECIDHGMGHWAVKIVKNGVRYEARSEIOVLEHLNSTDPNSVFCVQM 189  
Db 170 GEGTFRVQVQVDRRGARVAKIKVKEKAEARLEINVLKINEKDPDKNKLQVM 229  
QY 190 LEWFDHGHVCIIVFELLGLSTYDFIKENSFLPFQIDHIROMAYQICQSFILHNKLTH 249  
Db 230 FDFPDYHGMICSPFELLGLSTDFELKNNLYLPYHQVRRHAFOLCQAVKFLHDKLTH 289  
QY 250 DLKPNILFVKSIVVYKNSKMRDERTLKNITDKIVDFGATYDDEHSTLVSTRHYRA 309  
Db 290 DLKPNILFVNSDYELTYNLEKRDERSVKSTAVRVDFGATFDEHSTLVSTRHYRA 349  
QY 310 PEVILALGWSQPCDWSIGCIIIEYLGFTVQTHDSKEHLAMMERILGPIQHMIOKTR 369  
Db 350 PEVILELGSQPCDWSIGCIIIEYLGFTVQTHDNREHLAMMERILGPIPSMIRKTR 409  
QY 370 KRKYPHNLQDWEHSSAGRYVRRCKPLKEPMLCHDBEHEKLFDLVRRMLEYDPTQIT 429  
Db 410 KQKYFGRGLDWDENTSAGRYVRENCKPLRLRYLTSEAEHQLFDLIESMLEYEPKELT 469  
QY 430 LDEALQHPFFDLK 443  
Db 470 LGEALQHPFFARLR 483

RESULT 5  
S53639  
protein kinase clk3 (EC 2.7.1.-) - human  
C:Species: Homo sapiens (man)  
C:Date: 28-Oct-1995 #sequence\_revision 07-Feb-1997 #text\_change 09-Jul-2004  
C:Accession: S53639; S71040  
R:Hanes, J.; von der Kammer, H.; Klaudiny, J.; Scheit, K.H.  
J. Mol. Biol. 244, 665-672, 1994  
A:Title: Characterization by cDNA cloning of two new human protein kinases. Evidence by submitted to the EMBL Data Library, January 1995  
A:Reference number: S53637; MUID:95082033; PMID:7990150  
A:Accession: S53639  
A:Molecule type: mRNA  
A:Residues: 1-490 <HAN>  
A:Cross-references: UNIPROT:P49761; GB:L29217  
R:Hanes, J.J.; der Kammer, H.; Klaudiny, J.J.; Scheit, K.H.  
submitted to the EMBL Data Library, January 1995  
A:Reference number: S71040  
A:Accession: S71040  
A:Molecule type: mRNA  
A:Residues: 1-131, "TG", 134-490 <HAW>  
A:Cross-references: EMBL:L29217; NID:G632971; PIDN:AAA61484.1; PID:G632972

C:Superfamily: human protein kinase clk1; protein kinase homology  
C:Keywords: alternative splicing; phosphotransferase  
F:154-433/Domain: protein kinase homology <KIN>

Query Match 54.8%; Score 1321; DB 2; Length 490;  
Best Local Similarity 55.3%; Pred. No. 2e-56;  
Matches 251; Conservative 60; Mismatches 101; Indels 42; Gaps 6;

QY 4 PLEASHSVEEDTHPSHYLEARSLENERDYRRRYVDEYRNDYCEGYVP-----R 51  
Db 43 PPRSRSRSHDRLP-----YQRRYRRRSDSYR---CEERSPSFGEDYDGPSRSR 90  
QY 52 HYHRDIESG-YRI-----HCSKSVRRSRSPKRNHCHSCSHQSRKSHRRKRKRSIED 105  
Db 91 HRRSRERGYPYTRKHAHCHK-----RTRSCSSASRSRQSSKRSRSVED 138  
QY 106 DEEGHLICQSGDVLARVEIVDTLGEAGFKVVECIDHGMGHWAVKIVKNGVRYEAR 165  
Db 139 DKEGHLVCRIGDMLQERYEIVGNLGEFTGKVECIDHARKSQVALKILRNVGKTRAA 198  
QY 166 RSEIQVLEHLNSTDPNSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPFQ 225  
Db 199 RLEINVLKIKKEKKNKFLCVLMSDFNFGHMCIAFELLGKNTFEFLKNNFPQYPLP 258  
QY 226 HIROMAYQICQSFILHNKLTHDLKPNILFVKSIVVYKNSKMRDERTLKNITDKIV 285  
Db 259 HVRHMAVQLCHALRFLHENQLTDLKPNILFVNSEPETLYNHEKSCSEKSVKNTSIRV 318  
QY 286 VDFGSATYDDEHSTLVSTRHYRAPEVILALGWSQPCDWSIGCIIIEYLGFTVQTHD 345  
Db 319 ADFGSATFDEHHTTVIATRHYPPEVILELGSQPCDWSIGCIIIEYLGFTVQTHD 378  
QY 346 SKEHLAMMERILGPIQHMIOKTRKRYFHHNQDWDHSHSAGRYVRRCKPLKEFMLCH 405  
Db 379 NREHLVMEKILGPIPSMIRKTRKQKYFKYKGLVWDENSDGRVYKENCPLKSYMLQD 438  
QY 406 DEEHEKLFDLVRRMLEYDPTQITLDEALQHPFF 439  
Db 439 SLEHVQLFDLRRMLEFDPQAQRITLAEALLHPFF 472

RESULT 6  
S70352  
protein kinase clk3 (EC 2.7.1.-), testis-specific - rat  
N:Alternate names: LAMMER kinase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004  
C:Accession: S70352  
R:Becker, W.; Kentrup, H.; Heukelbach, J.; Joost, H.G.  
Biochim. Biophys. Acta 1312, 63-67, 1996  
A:Title: cDNA cloning and characterization of rat Clk3, a LAMMER kinase predominately expressed in testis  
A:Reference number: S70352; MUID:96271481; PMID:8679717  
A:Accession: S70352  
A:Molecule type: mRNA  
A:Residues: 1-490 <BEC>  
A:Cross-references: UNIPROT:Q63117; EMBL:X94351; NID:g1149536; PIDN:CAA64076.1; PID:g1149536  
A:Genetics:  
A:Gene: clk3  
C:Superfamily: human protein kinase clk1; protein kinase homology  
C:Keywords: phosphotransferase; protein kinase  
F:154-433/Domain: protein kinase homology <KIN>

Query Match 54.6%; Score 1317; DB 2; Length 490;  
Best Local Similarity 55.1%; Pred. No. 3.2e-56;  
Matches 250; Conservative 61; Mismatches 101; Indels 42; Gaps 6;

QY 4 PLEASHSVEEDTHPSHYLEARSLENERDYRRRYVDEYRNDYCEGYVP-----R 51  
Db 43 PPRSRSRSHDRIP-----YQRRYRRRSDSYR---CEERSPSFGEDYDGPSRSR 90  
QY 52 HYHRDIESG-YRI-----HCSKSVRRSRSPKRNHCHSCSHQSRKSHRRKRKRSIED 105  
Db 91 HRRSRERGYPYTRKHAHCHK-----RTRSCSSASRSRQSSKRSRSVED 138



[illegible]

```

RESULT 11
T04125
protein kinase PK12 (EC 2.7.1.1), ethylene-induced - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 16-Aug-2004
C:Accession: T04125
R:Sessa, G.; Raz, V.; Savaldi, S.; Fluhr, R.
Plant Cell 8, 2223-2234, 1996
A:Title: PK12, a plant dual-specificity protein kinase of the LAMMER family,
A:Reference number: Z15228; MUID:97143872; PMID:8989879
A:Accession: T04125
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-431 <SES>
A:Cross-references: UNIPROT:O45967; EMBL:U73937; NID:g2911279; PIDN:AAC04324.1
A:Experimental source: cultivar Samsun NN
C:Genetics:
A:Gene: PK12
C:Superfamily: protein kinase homology
C:Keywords: ATP; phosphotransferase; protein kinase
F:94-381/Domain: protein kinase homology <KIN>

```

| Query Match           | 34.8%        | Score 839;  | DB 2;           | Length 431; |
|-----------------------|--------------|---|-----------------|-------------|
| Best Local Similarity | 39.9%        | Pred. No. 1.7e-33;  |                 |             |
| Matches 175;          | Conservative | 80;   | Mismatches 130; | Indels 54;  |
| Gaps                  | 10           |   |                 |             |
| QY                    | 29           | RDYDRDYVDEYRNDYCEGVPRHYHRTDESGRVYHCSKSSVRSR                   | -----           | 73          |
| Db                    | 11           | RPYMDRPPKPRLLDMDPSHTPK-AOSGIYVGOEYVGNSSSYVHSRLLPDHPDLSLVKGLAQ | 69              |             |
| QY                    | 74           | RSSPKRKNRHCSSHQSRSSKSHRRKRSRSTEDDEGHLICQSGDVLRARVETDVLGEGA    | 133             |             |
| Db                    | 70           | KGSPPRR-----DDKQGHYMFELGENUTTRYKLLKKIGET                      | 106             |             |
| QY                    | 134          | FGKYVECIDHGMDBMHVAVKIVKNVGRYREAAARSETQVLEHLNSTDPNSVFRCVQMLEWF | 193             |             |
| Db                    | 107          | FGQVLECMWDREQGF-VAIKIIRSIKKYREAAVVEVDVLQLLGRYDRGCT-RCVQLRNWF  | 164             |             |
| QY                    | 194          | DHHGHCVIFELGLSTYDFIKNSFLPQIDHIROMAYQICQSNFLHNNKLTHTDLP        | 253             |             |
| Db                    | 165          | DYRNHI CLVPEKLGPSLDFELRNKSYRFPVDLREIGRQLLECAVAFHMDMLRIHTDLP   | 224             |             |
| QY                    | 254          | ENILFVKSDDV---VKYNSKMKRD-----ERTIKNTDIKVDPFGSATYDDEHHSITLAVSR | 305             |             |

|     |    |   |     |
|-----|----|---|-----|
| 225 | Db | ENILFVSADYIKVPDYKGTGTPWSHRDRSFKRLPFKSSAIKVIDFGSTAYERPDPHNTIVSTR | 284 |
| 306 | Qy | HYRAPEVILAGHSQPCDWSIGCIIIEYVLGFTVPTQTHDSKEHLAMMERIIGPIQHWI      | 365 |
| 285 | Db | HYRAPEVILGLWSYPCDLMSVGCIIELCSGEALFQTHENLEHLAMMERVLGPLPSQWL      | 344 |
| 366 | Qy | QKTRK-RKYFHHNOLDWDHSHSAGRYVRRCK-PLKEFMLC-DEEHEKLFDLVRRL         | 420 |
| 345 | Db | KVRDRHAETVRRGLDWPPEGATSRESIKSVMKLPRQLQNLVHQVHDSAGDLIDLQGLL      | 404 |
| 421 | Qy | EYDPTQRITLDEALQHPPF   | 439 |



Job time : 44 secs

```

Db      229 DDDGGHYKVPNSKFNRYTVVRLGHGTFGKVIQCYDQS-TGRHCAIKVTRAIKPYREA 287
QY      165 ARSEIOVLEHLNSTDPNSVRCVQMLEWFDHGHVCIVFELLGLSTYDFIKENSFLPFOI 224
Db      288 SLIELRVLTIAHSDPTNENKCIQLRDYFDYRKHCIVTDLFGWSVDFLKNNYIPFPL 347
QY      225 DHIROMAYQICQSINFLEHKLTHTDLPENILFKVSDYVVVKYNSKMKEDERTLKNNTDIK 284
Db      348 KHIQMLSQQLFKSVAFHLSGLVHTDLKPEVLLVSNASRTIRLPYRNYSQVLNSCEIR 407
QY      285 VVDFGSATYDDDEHSTLSTVTRHYRAPEVILALGWSQPCDWSIGCILLIYYILGFTVFQTH 344
Db      408 LIDFGSATFEDEHSSVSVTRHYRAPEIILGLGWSYPCDWSIGCILLIYELFTGQALFQTH 467
QY      345 DSKEHLAMMERILGPPOHMI---OKTRKRKYFHHNQLDWDHSSAGRYVR--RRCKPLK 399
Db      468 EDSEHLCKMEKILGPPDRNMISRRSRTSQRFFKSGKVRPLSNTPKKSINYLQSLQTLLE 527
QY      400 EFWLCHDEHEKLFDLVRMLEYDPTQRTITLDEALQHPFF 439
Db      528 QIFAVSSPEVALLDLKKVYVYDPKRRITAKEALWHPFF 567

RESULT 15
T38052
probable protein kinase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T38052
R.Lyle, G.; Church, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1995
A:Reference number: Z21765
A:Accession: T38052
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-690 <LYE>
A:Cross-references: EMBL:Z69239; PIDN:CAA93220.1; GSPDB:GN00066; SPDB:SPAC1D4.11c
A:Experimental source: strain 972h; cosmid cid4
C:Genetics:
A:Gene: SPDB:SPAC1D4.11c
A:Map position: 1
C:Superfamily: yeast probable protein kinase KNS1; protein kinase homology

Query Match      31.8%; Score 767; DB 2; Length 690;
Best Local Similarity 41.5%; Pred. No. 7.1e-30;
Matches 141; Conservative 78; Mismatches 115; Indels 6; Gaps 3;

QY      105 DDEGHLICQSGDVLRYEIVDTLGEAGKVEICIDHGMGMVAVKIVKNGRYREA 164
Db      344 DDDGGHYKVPNSKFNRYTVVRLGHGTFGKVIQCYDQS-TGRHCAIKVTRAIKPYREA 402
QY      165 ARSEIOVLEHLNSTDPNSVRCVQMLEWFDHGHVCIVFELLGLSTYDFIKENSFLPFOI 224
Db      403 SLIELRVLTIAHSDPTNENKCIQLRDYFDYRKHCIVTDLFGWSVDFLKNNYIPFPL 462
QY      225 DHIROMAYQICQSINFLEHKLTHTDLPENILFKVSDYVVVKYNSKMKEDERTLKNNTDIK 284
Db      463 KHIQMLSQQLFKSVAFHLSGLVHTDLKPEVLLVSNASRTIRLPYRNYSQVLNSCEIR 522
QY      285 VVDFGSATYDDDEHSTLSTVTRHYRAPEVILALGWSQPCDWSIGCILLIYYILGFTVFQTH 344
Db      523 LIDFGSATFEDEHSSVSVTRHYRAPEIILGLGWSYPCDWSIGCILLIYELFTGQALFQTH 582
QY      345 DSKEHLAMMERILGPPOHMI---OKTRKRKYFHHNQLDWDHSSAGRYVR--RRCKPLK 399
Db      583 EDSEHLCKMEKILGPPDRNMISRRSRTSQRFFKSGKVRPLSNTPKKSINYLQSLQTLLE 642
QY      400 EFWLCHDEHEKLFDLVRMLEYDPTQRTITLDEALQHPFF 439
Db      643 QIFAVSSPEVALLDLKKVYVYDPKRRITAKEALWHPFF 682

```

Search completed: March 13, 2005, 00:08:47

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2005, 00:05:12 ; Search time 29 Seconds  
(without alignments)  
726.868 Million cell updates/sec

Title: US-10-801-671-2  
Perfect score: 2410  
Sequence: 1 MCIPLEASHSVEDTHPSHY.....QRITLDEALQHPPDLLKKK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 190900 seqs, 47368951 residues

Total number of hits satisfying chosen parameters: 190900

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/paa/pct\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US11\_NEW\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query  | Score | Match | Length | DB                  | ID | Description        |
|------------|--------|-------|-------|--------|---------------------|----|--------------------|
| 1          | 1692   | 70.2  | 374   | 1      | PCT-US04-30360-133  |    | Sequence 133, App  |
| 2          | 1314.5 | 54.5  | 490   | 6      | US-10-450-763-53713 |    | Sequence 53713, A  |
| 3          | 1310   | 54.4  | 490   | 6      | US-10-878-630A-2    |    | Sequence 2, Appli  |
| 4          | 473.5  | 19.6  | 421   | 1      | PCT-US04-30360-134  |    | Sequence 134, App  |
| 5          | 472.5  | 19.6  | 763   | 1      | PCT-US04-42360-2024 |    | Sequence 2024, Ap  |
| 6          | 451.5  | 18.7  | 429   | 1      | PCT-US04-30360-135  |    | Sequence 135, App  |
| 7          | 420    | 17.4  | 1043  | 7      | US-11-050-926-258   |    | Sequence 258, App  |
| 8          | 419.5  | 17.4  | 371   | 1      | PCT-US04-30360-136  |    | Sequence 136, App  |
| 9          | 380    | 15.8  | 373   | 1      | PCT-US04-30360-35   |    | Sequence 35, Appl  |
| 10         | 337    | 14.0  | 703   | 6      | US-10-450-763-46352 |    | Sequence 46352, A  |
| 11         | 335.5  | 13.9  | 1030  | 1      | PCT-US05-03526-8    |    | Sequence 8, Appli  |
| 12         | 326.5  | 13.5  | 358   | 1      | PCT-US04-30360-60   |    | Sequence 60, Appl  |
| 13         | 326.5  | 13.5  | 359   | 7      | US-11-021-951-187   |    | Sequence 187, App  |
| 14         | 326.5  | 13.5  | 480   | 6      | US-10-450-763-44291 |    | Sequence 44291, A  |
| 15         | 323.5  | 13.4  | 359   | 1      | PCT-US04-30360-17   |    | Sequence 17, Appl  |
| 16         | 322.5  | 13.4  | 360   | 1      | PCT-US04-42360-1761 |    | Sequence 1761, Ap  |
| 17         | 322.5  | 13.4  | 360   | 1      | PCT-US04-42360-2389 |    | Sequence 2389, App |
| 18         | 322    | 13.4  | 355   | 1      | PCT-US04-30360-19   |    | Sequence 19, Appl  |
| 19         | 322    | 13.4  | 358   | 1      | PCT-US04-42360-2165 |    | Sequence 2165, Ap  |
| 20         | 321    | 13.3  | 360   | 1      | PCT-US04-30360-56   |    | Sequence 56, Appl  |
| 21         | 321    | 13.3  | 360   | 1      | PCT-US04-42360-2167 |    | Sequence 2167, Ap  |
| 22         | 319.5  | 13.3  | 326   | 1      | PCT-US04-30360-15   |    | Sequence 15, Appl  |
| 23         | 319.5  | 13.3  | 326   | 1      | PCT-US04-30360-52   |    | Sequence 52, Appl  |
| 24         | 316    | 13.1  | 304   | 7      | US-11-033-515-13    |    | Sequence 13, Appl  |
| 25         | 316    | 13.1  | 304   | 7      | US-11-033-515-23    |    | Sequence 23, Appl  |

|    |       |      |     |   |                     |  |                   |
|----|-------|------|-----|---|---------------------|--|-------------------|
| 26 | 313   | 13.0 | 367 | 1 | PCT-US04-42360-345  |  | Sequence 345, App |
| 27 | 313   | 13.0 | 367 | 1 | PCT-US04-42360-354  |  | Sequence 354, App |
| 28 | 312   | 12.9 | 367 | 1 | PCT-US04-30360-16   |  | Sequence 16, Appl |
| 29 | 312   | 12.9 | 396 | 1 | PCT-US04-30360-47   |  | Sequence 47, Appl |
| 30 | 312   | 12.9 | 411 | 1 | PCT-US04-30360-18   |  | Sequence 18, Appl |
| 31 | 312   | 12.9 | 431 | 6 | US-10-450-763-41854 |  | Sequence 41854, A |
| 32 | 310.5 | 12.9 | 335 | 1 | PCT-US04-30360-112  |  | Sequence 112, App |
| 33 | 309   | 12.8 | 367 | 1 | PCT-US04-42360-343  |  | Sequence 343, App |
| 34 | 309   | 12.8 | 367 | 1 | PCT-US04-42360-352  |  | Sequence 352, App |
| 35 | 303.5 | 12.6 | 356 | 7 | US-11-021-951-188   |  | Sequence 188, App |
| 36 | 301   | 12.5 | 334 | 1 | PCT-US04-30360-111  |  | Sequence 111, App |
| 37 | 299.5 | 12.4 | 338 | 1 | PCT-US04-30360-97   |  | Sequence 97, Appl |
| 38 | 292   | 12.1 | 527 | 7 | US-11-033-515-2     |  | Sequence 2, Appli |
| 39 | 290.5 | 12.1 | 368 | 8 | US-60-643-717-14615 |  | Sequence 14615, A |
| 40 | 288   | 12.0 | 307 | 1 | PCT-US04-42360-1763 |  | Sequence 1763, Ap |
| 41 | 288   | 12.0 | 307 | 1 | PCT-US04-42360-2391 |  | Sequence 2391, Ap |
| 42 | 286.5 | 11.9 | 337 | 8 | US-60-643-717-10080 |  | Sequence 10080, A |
| 43 | 286   | 11.9 | 390 | 1 | PCT-US04-30360-98   |  | Sequence 98, Appl |
| 44 | 284   | 11.8 | 303 | 7 | US-11-033-515-12    |  | Sequence 12, Appl |
| 45 | 284   | 11.8 | 303 | 7 | US-11-033-515-22    |  | Sequence 22, Appl |

ALIGNMENTS

RESULT 1  
PCT-US04-30360-133  
; Sequence 133, Application PC/TUS0430360  
; GENERAL INFORMATION:  
; APPLICANT: PLEXIKON, INC.  
; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT  
; FILE REFERENCE: 039363-1703  
; CURRENT APPLICATION NUMBER: PCT/US04/30360  
; PRIOR FILING DATE: 2004-09-15  
; PRIOR APPLICATION NUMBER: 60/503,277  
; PRIOR FILING DATE: 2003-09-15  
; NUMBER OF SEQ ID NOS: 167  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 133  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Mammalian  
; OTHER INFORMATION: protein sequence  
PCT-US04-30360-133

|                       |       |  |           |            |     |        |      |
|-----------------------|-------|--|-----------|------------|-----|--------|------|
| Query Match           | 70.2% | Score  | 1692;     | DB         | 1;  | Length | 374; |
| Best Local Similarity | 84.7% | Pred. No.  | 1.4e-101; |            |     |        |      |
| Matches               | 315;  | Conservative   | 24;       | Mismatches | 31; | Indels | 2;   |
| QY                    | 75    | SSPKRK-RNRHCSSH-QSRKSHRRKRSRSTEDDEGHLCQSGDVLRYEIVDTLGG       | 132       |            |     |        |      |
| Db                    | 1     | SSYKSKRIHHSTSHRRSHGSHRRKTRSVDEDEGHLCQSGDVLRYEIVDTLGG         | 60        |            |     |        |      |
| QY                    | 133   | AFGKVECDIHGMDGMHVAVKIVKNGRYREAAARSEIQVLEHLNSTDPNSVRCVQMLEW   | 192       |            |     |        |      |
| Db                    | 61    | AFGKVECDIHGMDGMHVAVKIVKNGRYREAAARSEIQVLEHLNSTDPNSVRCVQMLEW   | 120       |            |     |        |      |
| QY                    | 193   | PDHGHVCIIVPELLGLSTYDFIKENSLFPFQIDHIRMAYQICOSINFLHKNLTHTLK    | 252       |            |     |        |      |
| Db                    | 121   | FEHGHVCIIVPELLGLSTYDFIKENSLFPFQIDHIRMAYQICOSINFLHKNLTHTLK    | 180       |            |     |        |      |
| QY                    | 253   | PENILFVKSDVVKVYKNSKMRDERTLKNQTDIKVDFGSAATYDDHSHSLVSTRHYRAPEV | 312       |            |     |        |      |
| Db                    | 181   | PENILFVQSDTTEAYNPKIKRDETLNPDIKVDFGSAATYDDHSHSLVSTRHYRAPEV    | 240       |            |     |        |      |
| QY                    | 313   | ILALGWQPCDWSIGCIIIEYILGFTVFQTHDSKEHLAMMERILGPIPOHMIQKTRK     | 372       |            |     |        |      |
| Db                    | 241   | ILALGWQPCDWSIGCIIIEYILGFTVFQTHDSKEHLAMMERILGPIPOHMIQKTRK     | 300       |            |     |        |      |
| QY                    | 373   | YFHNQLDWDHSHSAGRYVRRCPKLFKEFMLCHDEEHEKLFDLVRMLLEYDPTQRTILDE  | 432       |            |     |        |      |

Db 301 YPHDRDLWDHSHSAGRYVSRACKFLKBFMLSQDVEHERLFDLIQKMLEYDPAKRITLRE 360

QY 433 ALQHPFFDLKK 444

Db 361 ALKHPFFDLKK 372

RESULT 2

US-10-450-763-53713  
; Sequence 53713, Application US/10450763  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 53713  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (271)..(290)  
; OTHER INFORMATION: TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE domain identified  
; OTHER INFORMATION: by eMATRIX, accession number PR00109B, p-value=7.559e-10, raw score of 12.27  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (155)..(470)  
; OTHER INFORMATION: Eukaryotic protein kinase domain identified by Pfam,  
; OTHER INFORMATION: accession name pkinase, E-value=4.8e-69, Pfam score of 242.8  
US-10-450-763-53713

Query Match 54.5%; Score 1314.5; DB 6; Length 490;

Best Local Similarity 59.8%; Pred. No. 2.8e-77;  
Matches 251; Conservative 59; Mismatches 89; Indels 21; Gaps 7;

QY 34 RRYVDE---YRNDYCEGVVPR-HYHRDIESGVRHCSKSSVSRSSPKRRK-----NR 83

Db 64 RYHSSEGRSGSYCEHRSRKHQRSS-----WSSSSDRTRR-----RRRDSYHVR 114

QY 84 HCSSHQSRKS-HRRKRSIEDDEGHLICQSGDVLRYEIVDTLGEAGFGKVECID 142

Db 115 RCSRFTFSRSSQHSRKAQSVEDDTEGHLIYHVGDLQERYEIVSTLKGTFGRVVCVD 174

QY 143 HGMDGMHVAIKVNGVRYRRAARSEIOVLEHLNSTDPNSVRCVQMLEWFDHGHVCIV 202

Db 175 HRRRGARVALKIIKNVEKYKEAARLEIKVLEKINEKDPGKNL-CVQMPDFDYGHCMS 233

QY 203 FELLGLSTYDFIKENSLPFOIDHRIQWYQICQSGINFLHNKLTHTDLKPENILFVKSD 262

Db 234 LELLGLSTFDLKNHNLPIYHQVHNASQCAVQKFLHDNKLTHDLKPENILFVNSD 293

QY 263 YVVKYNSRWKRDERTLKNLTDIKVVDVFGSATYDDEHSTLVSTRHYRAPEVILALGWSQPC 322

Db 294 YELTYNLEKKHERSVKSTAVRGVDFGSATFDEHSHSTIVSTRHYRAPEVILELGSQPC 353

QY 323 DVWSIGCILIIYYLGTFTQTHDSKEHLAMMERILGPQPMIOKTRKRYFHHNQDWD 382

Db 354 DVWSIGCIIFYYVGTFTQTHDNQHLATMERILGPISPMIRKTRKQKYFYRGLDWD 413

QY 383 EHSAGVRRRCRCKPLKEFMLCHDEHEKLFDLVRMLEYDPTORTITLDEALQHPFFDL 442

Db 414 ENTSGAVRRENCPLROYLTISEAEDHQLFDLIESMLEYEPQORLTIGEALQHPFFSRL 473

RESULT 3

US-10-878-630A-2  
; Sequence 2, Application US/10878630A  
; GENERAL INFORMATION:  
; APPLICANT: Wisotzkey, Robert G.  
; TITLE OF INVENTION: CDC28-Like protein kinase CLK3 disruptions, compositions and  
; FILE REFERENCE: R-3217/40338.0241USU1  
; CURRENT APPLICATION NUMBER: US/10/878,630A  
; CURRENT FILING DATE: 2004-06-28  
; PRIOR APPLICATION NUMBER: US 60/483,017  
; PRIOR FILING DATE: 2003-06-26  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 2  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-878-630A-2

Query Match 54.4%; Score 1310; DB 6; Length 490;

Best Local Similarity 55.6%; Pred. No. 5.4e-77;  
Matches 248; Conservative 62; Mismatches 110; Indels 26; Gaps 4;

QY 4 PLEASHSVEEDTHPSHYLEARSLNERDYDRRYVDEYRNDYCEGVVPRHYRHDIESGYRI 63

Db 43 PPRSRSRSHDRIP-----YQRRVREHRDSDTYR---CEERSPSFG---EDCYGS 86

QY 64 HCSSKSVSRSSPKR-----KRNHCSSHQSRKSHRRKRSRSEDDEGHLIC 113

Db 87 SRSRHRRSRERAPYRTRKHAHCHKRTSCSSASSRSQSSKRSRSDVDDKEGHLVC 146

QY 114 OSGDVLRYEIVDTLGEAGFGKVECIDHGMDGMHVAIKVNGVRYRRAARSEIOVLE 173

Db 147 RIGDWLQERYEIVGNLGEFTGKVVVECLDHARGSQVALKIIRNGKTRRARLEINLVK 206

QY 174 HLNSTDPNSVRCVQMLEWFDHGHVCIVFELLGLSTYDFIKENSLPFOIDHRIQWYQ 233

Db 207 KIKEDKENKFLCVLMSDFNFGHMCIAPELLKNTFEFLKNNFQYPLPHVRHMAVQ 266

QY 234 ICQSGINFLHNKLTHTDLKPENILFVKSDYVVKYNSRWKRDERTLKNLTDIKVVDVFGSATY 293

Db 267 LCHARFLHENQLTHTDLKPENILFVNSPEFTLYNEHKSCSEKSVKNTSIRVADFGSATF 326

QY 294 DDEHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYYLGTFTQTHDSKEHLAMM 353

Db 327 DHEHTTIVATRHYPPEVILELGAQPCDVWSIGCILFEYRGTFTLQTHENREHLVMM 386

QY 354 ERILGPIPOHMIQTRKRYFHHNQDWDHSSAGRYVRRCKPLKEFMLCHDEHEKLF 413

Db 387 EKILGPIPSHMIHTRKQKYFGKGLVWDENSSDGRVVKENCKPLKSYMLQDSLEHVQLF 446

QY 414 DLVRMLEYDPTORTITLDEALQHPFF 439

Db 447 DLMRMLEFDPAQRITLAEALLHPFF 472

RESULT 4

PCT-US04-30360-134  
; Sequence 134, Application PC/TUS0430360  
; GENERAL INFORMATION:  
; APPLICANT: PLEXIKON, INC.  
; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT  
; FILE REFERENCE: 039363-1703  
; CURRENT APPLICATION NUMBER: PCT/US04/30360  
; CURRENT FILING DATE: 2004-09-15  
; PRIOR APPLICATION NUMBER: 60/503,277  
; PRIOR FILING DATE: 2003-09-15  
; NUMBER OF SEQ ID NOS: 167  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 134  
; LENGTH: 421



```

QY 428 ITLDEALQHPF 438
Db 368 IPTIETLNHPF 378

RESULT 7
US-11-050-926-258
; Sequence 258, Application US/11050926
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHWANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLAT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: Assessment, Prevention, and Therapy of Ovarian Cancer
; CURRENT APPLICATION NUMBER: MRI-030
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-926-258

Query Match 17.4%; Score 420; DB 7; Length 1043;
Best Local Similarity 26.2%; Pred. No. 1.2e-19;
Matches 126; Conservative 91; Mismatches 183; Indels 80; Gaps 16;

QY 1 MCIPLEASHVEETH-----PSHYLEARSINERDYYRRVDEYRNDYCEGYVRHYR 55
Db 597 MSVPSEPS-SQSSRTTRSPDDILERVAADVKEY-ERENVDTF----- 639
QY 56 DIEGYRTHCSKSSVRSRRSPKRR-----NRHCSHOSRSKSHRR--KRS 100
Db 640 --EASVRAKHLMTVEQNNGSQKLLAPDMFTSDDMFAAYFDSARLARAGICKDFKEN 697
QY 101 RSIED---DEGHLCQSGDVLRYARYEIVDTLGGAGFGKVCEDHGMGMHVAVKI 157
Db 698 PNLRDNWTDAEGYRVNIGEVLDKRYNYGYTGQGVFSNVVRARNARANQEVAVKI 757
QY 158 VGRYREAAARSEIQVLEHLNSTDPNSVFCVQMLEWFDHGHVCIVFELLGLSTYDFKE- 216
Db 758 NEMLMQKTGLKELEFLKLNADADPDDKFHCLFRHFYHKQHLCLVFEPLSNLREVLK 817

```

```

QY 217 NSFLPFOIDHIRMAYQICQSFNLFHNNKLTHTDKPENILFVKSDYVVKYNSKMRDER 276
Db 818 GKDVGLHIKAVRSYSQQLFLALKKRCNHLHADIKPDNLVNES----- 862
QY 277 TLKNTDIKVDVFGSATY--DDEHSTLSTVSTRHYRAPEVILALGWSQPCDVMSIGCILLEY 334
Db 863 ---KTILKLCDFGSASHVADNDITPYLVS-RFYRAPEIIGKSYDYDGIDMWSVGCTLYEL 918
QY 335 YLGFTVQTGTHDSKEHLAMMERILGPIQHMIOK--TRRKYFHHN-----QLDWEHSSAGR 389
Db 919 YTGKILFPKGTNNHMLKAMDLGKGMKPNKIRKGVFKDQDFDQNLNFWYIEVDKTEREK 978
QY 390 Y-VRRRCCKPLKEFM-----LCHDEBEH--KLFDLVRRLMEYDPTORTILDEALQHPF 438
Db 979 VTMSTINPTKDLLADLIGCORLPEDQKVKYHQLKDLLDQILMLDPKAKRISINQALQHP 1038

RESULT 8
PCT-US04-30360-136
; Sequence 136, Application PC/TUS0430360
; GENERAL INFORMATION:
; APPLICANT: PLEXIKON, INC.
; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT
; FILE REFERENCE: 039363-1703
; CURRENT APPLICATION NUMBER: PCT/US04/30360
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/503,277
; PRIOR FILING DATE: 2003-09-15
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 136
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Mammalian
; OTHER INFORMATION: protein sequence
PCT-US04-30360-136

Query Match 17.4%; Score 419.5; DB 1; Length 371;
Best Local Similarity 29.5%; Pred. No. 4.6e-20;
Matches 107; Conservative 73; Mismatches 142; Indels 41; Gaps 10;

QY 98 KRSRSED---DERGHLCQSGDVLRYARYEIVDTLGGAGFGKVCEDHGMGMHVAVKI 154
Db 23 KENPLRDNWTDAEGYRVNIGEVLDKRYNYGYTGQGVFSNVVRARNARANQEVAVKI 82
QY 155 VKNVGYREAAARSEIQVLEHLNSTDPNSVFCVQMLEWFDHGHVCIVFELLGLSTYDFI 214
Db 83 IRNNELMQKTGLKELEFLKLNADADPDDKFHCLFRHFYHKQHLCLVFEPLSNLREVL 142
QY 215 KE-NSFLPFOIDHIRMAYQICQSFNLFHNNKLTHTDKPENILFVKSDYVVKYNSKMR 273
Db 143 KKYGKDVGLHIKAVRSYSQQLFLALKKRCNHLHADIKPDNLVNES----- 190
QY 274 DERTLKNTDIKVDVFGSATY--DDEHSTLSTVSTRHYRAPEVILALGWSQPCDVMSIGCIL 331
Db 191 -----KTILKLCDFGSASHVADNDITPYLVS-RFYRAPEIIGKSYDYDGIDMWSVGCTL 243
QY 332 IEYLGFTVQTGTHDSKEHLAMMERILGPIQHMIOK--TRRKYFHHN-----QLDWEHSS 386
Db 244 YELGTGKILFPKGTNNHMLKAMDLGKGMKPNKIRKGVFKDQDFDQNLNFWYIEVDKYTE 303
QY 387 AGRY-VRRRCCKPLKEFM-----LCHDEBEH--KLFDLVRRLMEYDPTORTILDEALQ 435
Db 304 REKVTVMSTINPTKDLLADLIGCORLPEDQKVKYHQLKDLLDQILMLDPKAKRISINQALQ 363
QY 436 HPF 438
Db 364 HAF 366

```

## RESULT 9

PCT-US04-30360-35  
; Sequence 35, Application PC/TUS0430360  
; GENERAL INFORMATION:  
; APPLICANT: PLEXIKON, INC.  
; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT  
; FILE REFERENCE: 039363-1703  
; CURRENT APPLICATION NUMBER: PCT/US04/30360  
; CURRENT FILING DATE: 2004-09-15  
; PRIOR APPLICATION NUMBER: 60/503,277  
; PRIOR FILING DATE: 2003-09-15  
; NUMBER OF SEQ ID NOS: 167  
; SOFTWARE: PatentIn ver. 3.2  
; SEQ ID NO 35  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Mammalian  
; OTHER INFORMATION: protein sequence  
PCT-US04-30360-35

Query Match 15.8%; Score 380; DB 1; Length 373;  
Best Local Similarity 27.9%; Pred. No. 1.6e-17; Mismatches 126; Indels 52; Gaps 10;  
Matches 96; Conservative 70; Mismatches 126; Indels 52; Gaps 10;  
QY 121 ARYEIVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGVRYREARSEIQVLEHLNSTDP 180  
DB 19 ARYLVRKLGGHGFSTWLAQDM-VNTHVAMKIVRGDKVYTEAAEDIKLQRVNDADN 77  
QY 181 N-----SVFRCVQMLEWFDHGG-----HVCIVFELLGLSTYDFIKENSPLPFOIDHIROMA 231  
DB 78 TKEDSMGANHILKLDHFNHKGPNGVHVMVFEVLGNLLALIKKYEHGRGPIYVVKQIS 137  
QY 232 YQICOSINFLHN-KLTHTDLKPENILFVKSQVYVVKYNSKMKRDERTLKNTDIKVVDFGS 290  
DB 138 KOLLGLDYMRRCGIITHTDKPENLMEIVD-----SPENLIQ-----IKIADLGN 184  
QY 291 ATYDEHSTLSTVSTHYRAPEVILALGWSQPCDVMSIGCILIEYVLGFTVPQTHD----- 345  
DB 185 ACWDEHYTNSIQTFEYSPEVLLGAPWCGGADIWSTACILIFELITGDFEPDEGHSYT 244  
QY 346 -SKEHLAMMERILGPIPOHMTQKTRKYFHHNQLDWDEHSSAGRYVRRCKPKL----- 399  
DB 245 KDDHIAQIIBLLGSLPALLRNGKYTRTFNS-----RGLLRNISKLFWPLE 293  
QY 400 -----EFMLCHDEEHEKFLDLVRMLEYDPTQRTLDALQHPF 438  
DB 294 DVLTEKYKFSKDEAKE-ISDFSLPMLQLDPRKRDAGGLVNHWP 336

## RESULT 10

US-10-450-763-46352  
; Sequence 46352, Application US/10450763  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/549,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 46352  
; LENGTH: 703  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

; NAME/KEY: DOMAIN  
; LOCATION: (622).. (653)  
; OTHER INFORMATION: Protein kinases ATP-binding region proteins domain identified  
; OTHER INFORMATION: by eMATRIX, accession number BL00107A, p-value=6.586e-11, raw  
; OTHER INFORMATION: score of 18.39  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (507).. (674)  
; OTHER INFORMATION: Eukaryotic protein kinase domain identified by PFam,  
; OTHER INFORMATION: accession name pkinase, E-values=3.5e-27, PFam score of 101.9  
US-10-450-763-46352

Query Match 14.0%; Score 337; DB 6; Length 703;  
Best Local Similarity 32.7%; Pred. No. 1.7e-14;  
Matches 65; Conservative 47; Mismatches 69; Indels 18; Gaps 2;  
QY 98 KRSSIEDDEBGLICQSGDVLRYRIEIVDTLGEAGFGKVEICIDHGMGMHVAVKIVKN 157  
DB 481 KFSKTSFDDDEHGFKVLVHDHIAIYRYEVLFTIGKSGFGQVAKCLDHKNEL-VALKIIRN 539  
QY 158 VGYREARSEIQVLEHLNSTDPNSVFRVCVQMLEWFDHGHVCIIVFELLGLSTYDFIKEN 217  
DB 540 KKRPHQALMELKILEALRKDKDNTYVNVHMKDFFYFRNHFCITFELLGINLYELMKNN 599  
QY 218 SFLPFOIDHIROMAYQICQSNFLHKNLTHTDLKPENILFVKSQVYVVKYNSKMKRDERT 277  
DB 600 NFOGFSLSIVRRFTLSVLKCLQMLSVKIIHCDLKPENILVLYQK----- 644  
QY 278 LKNTDIKVVDFGSATYDDE 296  
DB 645 --QASVKVIDFGSSCYEHQ 661

## RESULT 11

PCT-US05-03526-8  
; Sequence 8, Application PC/TUS0503526  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: MBMS AS MODIFIERS OF BRANCHING MORPHOGENESIS AND METHODS OF USE  
; FILE REFERENCE: EX05-003C-PC  
; CURRENT APPLICATION NUMBER: PCT/US05/03526  
; CURRENT FILING DATE: 2005-01-27  
; PRIOR APPLICATION NUMBER: US60/539,835  
; PRIOR FILING DATE: 2004-01-28  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 8  
; LENGTH: 1030  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US05-03526-8

Query Match 13.9%; Score 335.5; DB 1; Length 1030;  
Best Local Similarity 26.1%; Pred. No. 3.1e-14;  
Matches 86; Conservative 74; Mismatches 125; Indels 45; Gaps 9;  
QY 116 GDVLRYRIEIVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGVRYREARSEIQVLEHL 175  
DB 7 GNVMN-KFEILGVVGEAGYGVVLC-RHKETHIIVAIKKPKDSENEVSKETTLRELKWL 64  
QY 176 NSTDPNSVFRVCVQMLEWFDHGHVCIIVFELLGLSTYDFIKE--NSFLPFOIDHIROMAYQ 233  
DB 65 RTLQENI---VELKEAFRRRGKLYLVFEYVEKNMELLEEMPNGVPP---EKVKSYYIQ 118  
QY 234 ICQSNFLHKNLTHTDLKPENILFVKSQVYVVKYNSKMKRDERTLKNTDIKVVDFGSATY 293  
DB 119 LKAIHWCHRNIDIVHRDILKPENLISHNDV-----LKLCDGFGFARN 159  
QY 294 ----DDEHSTLSTVSTHYRAPEVILALGWSQPCDVMSIGCILIEYVLGFTVPQTHDSKEH 349  
DB 160 LSEGNANYTEYVATRWYRSPELLLGAPYKGSVDWMSVGCILGELSDGQPLPFGSEIDQ 219  
QY 350 LAMMERILGPIPOHMTQKTRKYFHHNQLDWDEHSSAGRYVRRCKPKLKFMLCHDSEH 409

Db 220 LFTIQKVLGPSEQKLFYSNPRFHGLRFPVAVNHPQS--LERRYLGL-----N 267  
 QY 410 EKLFDLVRMLDYDTQRTITLDEALQHFF 439  
 Db 268 SVLLDMKNLLKLPADRYLTEQCLNHPTF 297

RESULT 12  
 PCT-US04-30360-60  
 ; Sequence 60, Application PC/TUS0430360  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PLEXIKON, INC.  
 ; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT  
 ; CURRENT APPLICATION NUMBER: PCT/US04/30360  
 ; CURRENT FILING DATE: 2004-09-15  
 ; PRIOR APPLICATION NUMBER: 60/503,277  
 ; PRIOR FILING DATE: 2003-09-15  
 ; NUMBER OF SEQ ID NOS: 167  
 ; SOFTWARE: PatentIn Ver. 3.2  
 ; SEQ ID NO 60  
 ; LENGTH: 358  
 ; TYPE: PRT  
 ; ORGANISM: Unknown Organism  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Unknown Organism: Mammalian  
 ; OTHER INFORMATION: protein sequence  
 PCT-US04-30360-60

Query Match 13.5%; Score 326.5; DB 1; Length 358;  
 Best Local Similarity 26.9%; Pred. No. 4.1e-14;  
 Matches 98; Conservative 64; Mismatches 131; Indels 71; Gaps 12;  
 QY 89 QRSKSHRRKRSRSTDEEGHLCQSGDVLRYEIVDTLGEAGFKVVECDHGM 148  
 Db 3 QERTFYRQELNKTWEVPE-----RYQNLSPVSGAYGVCVCAAFD-TKTGL 48

QY 149 HVAVK-----IVKNVGYREARSEIOVLEHL---NSTDPNSVFCVQMLEWFDHG 197  
 Db 49 RVAVKLSRPQSIHAKRTYR-----ELRLKMKHENVIGLLDVTTPARSLSEFN--- 100  
 QY 198 HVCIVFELLGLSTYDFIKENSFLPQIDHIRQAYICQSFNHLHKNLTHTDLPENIL 257  
 Db 101 DVYLVTWLMGADLNNIVKQCLTD---DHVQFLYQLRLGLKYIHSADI IHRDLKPSN-L 156  
 QY 258 FVKSQPCDVSIGCILIEYLGFTVQTHDSKEHLAMMERILGPIPOHMIQKTRKRYFHH 316  
 Db 157 AVNED-----CELKILDFGLARHTDDEMTGYVATRWYRAPEIMLNWM 198  
 QY 317 GWSQPCDVSIGCILIEYLGFTVQTHDSKEHLAMMERILGPIPOHMIQKTRKRYFHH 376  
 Db 199 HYNQTVDIWSVGCIMAEILLTGRTPFGTDHINQIQINMLTGTPPAYLINRMPSE----- 254  
 QY 377 NOLDWDEHSSAGRYVRRCK-PLKEFMLCHDEHEKFLDLVRRMLDYDPTQRTITLDEALQ 435  
 Db 255 -----ARNYQISLTQMPKMFANVFIGNPLAVDLLEKMLVLDSDKRITAAQALA 304  
 QY 436 HPFF 439  
 Db 305 HAYF 308

RESULT 13  
 US-11-021-951-187  
 ; Sequence 187, Application US/11021951  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HAUPTS, Ulrich  
 ; APPLICANT: KOLTERMANN, Andre  
 ; APPLICANT: SCHEIDIG, Andreas  
 ; APPLICANT: VOTSMELER, Christian  
 ; APPLICANT: Kettling, Ulrich  
 ; APPLICANT: COCO, Wayne Michael

; TITLE OF INVENTION: New Biological Entities And The Pharmaceutical  
 ; TITLE OF INVENTION: And Diagnostic Use Thereof  
 ; FILE REFERENCE: 04156.0002US  
 ; CURRENT APPLICATION NUMBER: US/11/021,951  
 ; CURRENT FILING DATE: 2004-12-22  
 ; PRIOR APPLICATION NUMBER: 10/872,198  
 ; PRIOR FILING DATE: 2004-06-18  
 ; PRIOR APPLICATION NUMBER: 60/543,518  
 ; PRIOR FILING DATE: 2004-02-11  
 ; PRIOR APPLICATION NUMBER: 60/524,960  
 ; PRIOR FILING DATE: 2003-11-25  
 ; PRIOR APPLICATION NUMBER: EP 04003058  
 ; PRIOR FILING DATE: 2004-02-11  
 ; PRIOR APPLICATION NUMBER: EP 03025871  
 ; PRIOR FILING DATE: 2003-11-11  
 ; PRIOR APPLICATION NUMBER: EP 03025851  
 ; PRIOR FILING DATE: 2003-11-10  
 ; PRIOR APPLICATION NUMBER: EP 03013819  
 ; PRIOR FILING DATE: 2003-06-18  
 ; NUMBER OF SEQ ID NOS: 191  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 187  
 ; LENGTH: 359  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-11-021-951-187

Query Match 13.5%; Score 326.5; DB 7; Length 359;  
 Best Local Similarity 26.9%; Pred. No. 4.1e-14;  
 Matches 98; Conservative 66; Mismatches 129; Indels 71; Gaps 12;  
 QY 89 QRSKSHRRKRSRSTDEEGHLCQSGDVLRYEIVDTLGEAGFKVVECDHGM 148  
 Db 2 QERTFYRQELNKTWEVPE-----RYQNLSPVSGAYGVCVCAAFD-TKTGL 47  
 QY 149 HVAVK-----IVKNVGYREARSEIOVLEHL---NSTDPNSVFCVQMLEWFDHG 197  
 Db 48 RVAVKLSRPQSIHAKRTYR-----ELRLKMKHENVIGLLDVTTPARSLSEFN--- 99  
 QY 198 HVCIVFELLGLSTYDFIKENSFLPQIDHIRQAYICQSFNHLHKNLTHTDLPENIL 257  
 Db 100 DVYLVTWLMGADLNNIVKQCLTD---DHVQFLYQLRLGLKYIHSADI IHRDLKPSN-L 155  
 QY 258 FVKSQPCDVSIGCILIEYLGFTVQTHDSKEHLAMMERILGPIPOHMIQKTRKRYFHH 316  
 Db 156 AVNED-----CELKILDFGLARHTDDEMTGYVATRWYRAPEIMLNWM 197  
 QY 317 GWSQPCDVSIGCILIEYLGFTVQTHDSKEHLAMMERILGPIPOHMIQKTRKRYFHH 376  
 Db 198 HYNQTVDIWSVGCIMAEILLTGRTPFGTDHIDQLKLILRLVGTGPAELLKKI----- 249  
 QY 377 NOLDWDEHSSAGRYVRRCK-PLKEFMLCHDEHEKFLDLVRRMLDYDPTQRTITLDEALQ 435  
 Db 250 -----SSARNYQISLTQMPKMFANVFIGNPLAVDLLEKMLVLDSDKRITAAQALA 303  
 QY 436 HPFF 439  
 Db 304 HAYF 307

RESULT 14  
 US-10-450-763-44291  
 ; Sequence 44291, Application US/10450763  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: 790CIP3/US  
 ; CURRENT APPLICATION NUMBER: US/10/450,763  
 ; CURRENT FILING DATE: 2003-06-11  
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: 09/540,217  
 ; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: 09/649,167  
 ; PRIOR FILING DATE: 2000-08-23  
 ; NUMBER OF SEQ ID NOS: 60736  
 ; SOFTWARE: Custom  
 ; SEQ ID NO 44291  
 ; LENGTH: 480  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (260)..(291)  
 ; OTHER INFORMATION: Protein kinases ATP-binding region proteins domain identified  
 ; OTHER INFORMATION: by eMATRIX, accession number BL00107A, p-value=4.375e-21, raw  
 ; OTHER INFORMATION: score of 18.39  
 ; US-10-450-763-44291

Query Match 13.5%; Score 326.5; DB 6; Length 480;  
 Best Local Similarity 26.9%; Pred No. 5.5e-14;  
 Matches 98; Conservative 64; Mismatches 131; Indels 71; Gaps 12;  
 Qy 89 QSRKSHRRKRSIEDDEGHLCQSGDVLRLARYEIVDTLGEAGFGKVVCEIDHGM DGM 148  
 Db 123 QERPTFYRQELNKTIEWPE-----RYQNLSPVSGGAYGVCVCAAFD-TKTGL 168  
 Qy 149 HVAVK-----IVKNGRYREARSEIQVLEHL---NSTDPNSVFRVQMLEWFDHGG 197  
 Db 169 RVAVKLSRPFQSIHAKRTYR-----ELRLKMKHENVIGLLDVFTPARSLEEFN--- 220  
 Qy 198 HVCIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSFNHLHNNKLTHTDLKPENIL 257  
 Db 221 DVLVTHLMGADLNNIVKCKQLTD---DHVQFLIYQILRLGLKYIHSADIHRDLKPSN-L 276  
 Qy 258 FVKSDDYVVKYNSKMKRDERLTKNNTDIKVDFGSAFYDDEHSTLSTVSTRHYRAPEVILA-L 316  
 Db 277 AVNED-----CELKILDFGLARHTDDEMTGYVATRWYRAPEIMLNWM 318  
 Qy 317 GWSQPCDWSIGCILIEYILGFTVFTQTHDSKEHLAMMERILGPPOHMIQTRKRYFHH 376  
 Db 319 HYNQTVDIWSVGCIMAEELLTGTLPFGTDHINQLQQLIMRLATGTPPAYLINRMP SHE----- 374  
 Qy 377 NQLDWEHSSAGRYVRRCK-PLKEFMLCHDEEKLFDLVRMLEYDPTQRIITLDEALQ 435  
 Db 375 -----ARNYIQSITQMPKMFANVFIGNPLAVDLLEKMLVLDSDKRITAAQAALA 424  
 Qy 436 HPFF 439  
 Db 425 HAYF 428

RESULT 15  
 PCT-US04-30360-17  
 ; Sequence 17, Application PC/TUS0430360  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEXIKON, INC.  
 ; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT  
 ; FILE REFERENCE: 039363-1703  
 ; CURRENT APPLICATION NUMBER: PCT/US04/30360  
 ; CURRENT FILING DATE: 2004-09-15  
 ; PRIOR APPLICATION NUMBER: 60/503,277  
 ; PRIOR FILING DATE: 2003-09-15  
 ; NUMBER OF SEQ ID NOS: 167  
 ; SOFTWARE: PatentIn Ver. 3.2  
 ; SEQ ID NO 17  
 ; LENGTH: 359  
 ; TYPE: PRT  
 ; ORGANISM: Unknown Organism  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Unknown Organism: Mammalian  
 ; OTHER INFORMATION: protein sequence  
 ; PCT-US04-30360-17

Query Match 13.4%; Score 323.5; DB 1; Length 359;  
 Best Local Similarity 26.9%; Pred. No. 6.4e-14;

Matches 98; Conservative 65; Mismatches 130; Indels 71; Gaps 12;  
 Qy 89 QSRKSHRRKRSIEDDEGHLCQSGDVLRLARYEIVDTLGEAGFGKVVCEIDHGM DGM 148  
 Db 2 QERPTFYRQELNKTIEWPE-----RYQNLSPVSGGAYGVCVCAAFD-TKTGH 47  
 Qy 149 HVAVK-----IVKNGRYREARSEIQVLEHL---NSTDPNSVFRVQMLEWFDHGG 197  
 Db 48 RVAVKLSRPFQSIHAKRTYR-----ELRLKMKHENVIGLLDVFTPARSLEEFN--- 99  
 Qy 198 HVCIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSFNHLHNNKLTHTDLKPENIL 257  
 Db 100 DVLVTHLMGADLNNIVKCKQLTD---DHVQFLIYQILRLGLKYIHSADIHRDLKPSN-L 155  
 Qy 258 FVKSDDYVVKYNSKMKRDERLTKNNTDIKVDFGSAFYDDEHSTLSTVSTRHYRAPEVILA-L 316  
 Db 156 AVNED-----CELKILDFGLARHTDDEMTGYVATRWYRAPEIMLNWM 197  
 Qy 317 GWSQPCDWSIGCILIEYILGFTVFTQTHDSKEHLAMMERILGPPOHMIQTRKRYFHH 376  
 Db 198 HYNQTVDIWSVGCIMAEELLTGTLPFGTDHINQLQQLIMRLATGTPPAYLINRMP SHE----- 249  
 Qy 377 NQLDWEHSSAGRYVRRCK-PLKEFMLCHDEEKLFDLVRMLEYDPTQRIITLDEALQ 435  
 Db 250 -----SSESARNYIQSITQMPKMFANVFIGNPLAVDLLEKMLVLDSDKRITAAQAALA 303  
 Qy 436 HPFF 439  
 Db 304 HAYF 307

Search completed: March 13, 2005, 00:18:18  
 Job time : 30 secs

**This Page Blank (uspto)**